

GenCore version 5.1.3
copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
 November 8, 2002, 19:54:37 ; Search time 103.462 seconds
 (without alignments)
 Run on: 11868.256 Million cell updates/sec

US-09-642-034-1
Title: 3461
perfect score: 1
sequence: 1 ctctgtgcgaattcggcagc.....aaattccttatatcagcttg 3461

```
scoring table: IDENTITY_NUC
                10 0 Gapext 1.0
```

177392727 residues

Searched: 320260 seqs, 17/39472, 2017-07-10 10:00:00
Chosen parameters: 640520

```

Total number of hits satisfying criteria: 1
Minimum DB seq length: 0
0000000000

```

post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first 45 summaries	

```

Database :
published_Applications_NA:*
1: /cgn2_6/ptodata1/pubpna/us07_PUBCOMB.seq:*
2: /cgn2_6/ptodata1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata1/pubpna/PCTUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata1/pubpna/us08_NEW_PUB.seq:*
7: /cgn2_6/ptodata1/pubpna/us08_PUBCOMB.seq:*
8: /cgn2_6/ptodata1/pubpna/us09_NEW_PUB.seq:*
9: /cgn2_6/ptodata1/pubpna/us09_PUBCOMB.seq:*
10: /cgn2_6/ptodata1/pubpna/us10_NEW_PUB.seq:*
11: /cgn2_6/ptodata1/pubpna/us10_PUBCOMB.seq:*
12: /cgn2_6/ptodata1/pubpna/us60_NEW_PUB.seq:*
13: /cgn2_6/ptodata1/pubpna/us60_PUBCOMB.seq:*
14: /cgn2_6/ptodata1/pubpna/us60_PUBCOMB.seq:*

```

14: /cgn2-9/p/locatc-7/r-1.

SUMMARIES

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	1141.4	33.0	1193	10	US-09-925-300-591	Sequence 591, App
2	404.8	11.7	483	10	US-09-867-701-2021	Sequence 2021, Ap
3	390.8	11.3	395	10	US-09-778-320-94	Sequence 94, Appl
C 2	390.8	11.3	395	10	US-09-910-689-94	Sequence 94, Appl
C 4	390.8	11.3	395	12	US-10-010-742-94	Sequence 94, Appl
C 5	390.8	11.3	395	12	US-10-010-742-94	Sequence 78, Appl
C 6	360	10.4	385	10	US-09-827-669-78	Sequence 14, Appl
7	319	9.2	5330	10	US-09-789-561-14	Sequence 360, App
C 8	240	6.9	241	10	US-09-607-805-360	Sequence 360, App
C 9	240	6.9	241	12	US-10-007-805-360	Sequence 162, App
C 10	165	4.8	177	10	US-09-759-143-162	Sequence 162, App
C 11	165	4.8	177	10	US-09-780-669-162	Sequence 162, App
C 12	165	4.8	177	10	US-09-630-606-162	Sequence 162, App
C 13	165	4.8	177	10	US-09-822-827-162	Sequence 162, App
C 14	165	4.8	177	10	US-09-115-453-162	Sequence 1649, Ap
C 15	127.6	3.7	4573	10	US-09-880-107-1649	Sequence 228, App
16	62.2	1.8	574	10	US-09-864-761-228	Sequence 17051, A
17	62.2	1.7	285	10	US-09-864-761-228	Sequence 64, Appl
C 18	59.8	1.7	342	10	US-09-737-178-65	Sequence 65, Appl
C 19	59.8	1.7	342	10	US-09-737-178-65	

C	20	59.8	1.7	351	10	US-09-737-178-60	Sequence 60, Appl
					10	US-09-737-178-63	Sequence 63, Appl
C	21	59.8	1.7	356	10	US-09-737-178-58	Sequence 58, Appl
C	22	59.8	1.7	358	10	US-09-737-178-66	Sequence 66, Appl
C	23	59.8	1.7	363	10	US-09-737-178-67	Sequence 67, Appl
C	24	59.8	1.7	363	10	US-09-737-178-67	Sequence 59, Appl
C	25	59.8	1.7	409	10	US-09-737-178-59	Sequence 61, Appl
C	26	59.8	1.7	410	10	US-09-737-178-61	Sequence 62, Appl
C	27	59.8	1.7	416	10	US-09-737-178-62	Sequence 42, Appl
				1271	10	US-09-737-178-5	Sequence 6, Appl
C	28	59.8	1.7	1271	10	US-09-737-178-42	Sequence 20421, A
C	29	59.8	1.7	1821	10	US-09-737-178-6	Sequence 3471, A
C	30	59.8	1.7	1635	10	US-09-864-761-20241	Sequence 7, Appl
C	31	59.4	1.7	1973	10	US-09-864-761-3471	Sequence 45, Appl
C	32	59.4	1.7	4223	10	US-09-737-178-7	Sequence 45, Appl
C	33	58.8	1.7	4223	10	US-09-737-178-45	Sequence 2321, Ap
C	34	59.8	1.7	468	10	US-09-864-761-2321	Sequence 1390, Ap
C	35	57.8	1.7	18596	10	US-09-864-107-1590	Sequence 119, App
C	36	57.2	1.7	18596	10	US-09-867-768A-119	Sequence 7674, App
C	37	57.2	1.7	18596	10	US-09-864-761-764	Sequence 20772, A
C	38	56.8	1.6	464	10	US-09-864-761-764	Sequence 19241, A
C	39	56.8	1.6	583	10	US-09-864-761-20772	Sequence 2513, A
C	40	56.8	1.6	1075	10	US-09-864-761-19241	Sequence 4012, Ap
C	41	56.8	1.6	1403	10	US-09-864-761-2513	Sequence 3936, Ap
C	42	56.8	1.6	1959	10	US-09-864-761-4012	Sequence 20699, A
C	43	56.2	1.6	401	10	US-09-864-761-3936	Sequence 1, Appl
C	44	56.2	1.6	446	10	US-09-864-761-20699	Sequence 42, Appl
C	45	55	1.6	492	10	US-09-737-178-1	

ALIGNMENTS

RESULT 1

```

RES001: 925-300-591
US-09-925-300-591, Application US/09925300
: Sequence 591,
: Patent No. US20020151681A1
: GENERAL INFORMATION:
: APPLICANT: Craig Rosen,
: APPLICANT: Steve Ruben,
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA101
: CURRENT APPLICATION NUMBER: US/09/925,300
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05988
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 1890
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 591
: LENGTH: 1193
: TYPE: DNA
: ORGANISM: Homo sapiens
ne-09-925-300-591

```

Query Match 33.0%; Score 1141.4; DB 10; Length 1193;
 Best Local Similarity 99.7%; Pred. No. 1.2e-266;
 Conservative 4; Mismatches 0; Indels 0; Gaps 0;

us-09-642-034-1.rnpb

Tue Nov 26 12:43:33 2002

QY 1989 TTTACTCAAGGCTTATCAAGTGGTTTAAAGTACTTCTGTGCTGTGTTCTGTCATGAGTTG 2048
|| |||||
Db 1194 TTCACGTGTGTCAGTTTCCAGGCATCAGCACCTCGGTGGCCATCCCTCTGTGAGGATTC 1253
|| |||||
QY 2049 CCTCATGAATTAGGTGACTTTGCTGTTCTACTAAAGGCTGGCATGACCGTTAAGCAGGCT 2108
|| |||||
Db 1254 CCACATCAGCTAGGAGACTTTGTCATCCCTGCTCAACGCTGGGATGAGCATCCAACAAGCT 1313
|| |||||
QY 2109 GTCCTTTATATGCAATGTGCAGCCATGCTGGGCTATCTTGGGAATGGCAACAGGAATTTTC 2168
|| |||||
Db 1314 CTCCTTCTCAACTTCCTTCTGCTGCTGCTACCTGGGCTGGCCCTTTGGCATCCTG 1373
|| |||||
QY 2169 ATTGCTCATATGCTGAAATGTTCTATGCTGGATATTTGCACTTACTGCTGGCTTATTC 2228
|| |||||
Db 1374 GCCGG-----CAGCCACTTCTCTGCCACTGGATTTTTCGCTAGCTGGAGGAATGTTTC 1427
|| |||||
QY 2229 ATCTATGTTGCTCTGTTGATATGTTACCTGAATG 2264
|| |||||
Db 1428 TTCTATATTTCTGCTGCTGATATGTTCCCTGAGATG 1463
|| |||||

Search completed: November 8, 2002, 22:05:19
Job time : 125.462 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 8, 2002, 16:42:31 ; Search time 6148.03 Seconds
(without alignments)
16383.279 Million cell updates/sec
Title: US-09-642-034-1
Perfect score: 3461
Sequence: 1 ctcgtgcgaattcgccacg.....aaattcttatcatcagcttg 3461
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.in.*
- 18: em.mu.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.ot.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_man.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3461	100.0	3461	6	AX465588	AX465588 Sequence
2	3391	98.0	3461	6	AX207205	AX207205 Sequence
3	2761.2	79.8	2776	6	AX207207	AX207207 Sequence
4	2716	78.5	2744	9	HSU41060	U41060 Homo sapien
5	2249	65.0	2404	6	176892	176892 Sequence 3
6	2134.2	61.7	2236	9	BC008317	BC008317 Homo sapi
7	1753.2	50.7	3287	10	AB071697	AB071697 Mus muscu
8	1486.8	43.0	1597	6	AX017261	AX017261 Sequence 1
9	1288	37.2	1310	6	176891	176891 Sequence 1
10	1137.8	32.9	221941	9	AC091060	AC091060 Homo sapi
11	1137.8	32.9	224788	2	AF001905	AF001905 Homo sapi
12	804	23.2	160170	2	AF001158	AF001158 Homo sapi
13	619.6	17.9	899	6	AX465590	AX465590 Sequence
14	549.4	15.9	551	6	AX207216	AX207216 Sequence
15	429.2	12.4	193437	2	AC126667	AC126667 Rattus no
16	390.8	11.3	395	6	AX156237	AX156237 Sequence
17	336.8	9.7	445	11	G24009	G24009 human STS W
18	319.2	9.2	337	6	AX393861	AX393861 Sequence
19	319	9.2	5231	9	AB033091	AB033091 Homo sapi
20	319	9.2	5322	6	AX405756	AX405756 Sequence
21	282.8	8.2	61103	2	AC131276	AC131276 Homo sapi
22	253.8	7.3	391	6	AX207209	AX207209 Sequence
23	240	6.9	241	6	AX303040	AX303040 Sequence
24	191.2	5.5	193437	2	AC126667	AC126667 Rattus no
25	165	4.8	177	6	AX106381	AX106381 Sequence
26	165	4.8	177	6	AX140672	AX140672 Sequence
27	165	4.8	177	6	AX200532	AX200532 Sequence
28	165	4.8	177	6	AX267188	AX267188 Sequence
29	144	4.2	52684	2	AC014796	AC014796 Drosophil
30	144	4.2	174712	3	AC010041	AC010041 Drosophil
31	144	4.2	195148	3	AC010019	AC010019 Drosophil
32	144	4.2	282267	3	AE003555	AE003555 Drosophil
33	129.2	3.7	2174	10	BC021530	BC021530 Mus muscu
34	128.4	3.7	2015	9	AB051127	AB051127 Macaca fa
35	127.6	3.7	4573	6	AX409002	AX409002 Sequence
36	127.6	3.7	4573	9	HUMORFG1P	D31887 Human mRNA
37	126	3.6	4610	6	AX078294	AX078294 Sequence
38	126	3.6	4698	6	AX281784	AX281784 Sequence
39	123.4	3.6	1383	6	AX118897	AX118897 Sequence
40	123.4	3.6	1803	9	AF193052	AF193052 Homo sapi
41	123.4	3.6	2604	9	AK027652	AK027652 Homo sapi
42	123.4	3.6	3169	9	BC012125	BC012125 Homo sapi
43	123.4	3.6	3246	9	AB040120	AB040120 Homo sapi
44	122.6	3.5	1895	9	BC027884	BC027884 Homo sapi
45	121.4	3.5	55577	2	AC014109	AC014109 Drosophil

ALIGNMENTS

RESULT 1	AX465588	AX465588	3461 bp	DNA	linear	PAT 16-JUL-2002
LOCUS	Sequence 1 from Patent WO0216939.					
DEFINITION	AX465588					
ACCESSION	AX465588.1	GI:21899891				
VERSION						
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.					
REFERENCE						
AUTHORS	Mack, D., Gish, K.C. and Wilson, K.E.					
TITLE	Methods of diagnosis of cancer and screening for cancer modulators					
JOURNAL	Patent: WO 0216939-A 1 28-FEB-2002;					

[illegible]

Db	1681	AGAGGTGATGATAGCTCATGCTCATCCACAGGAGTGTACATGAATATGTACCCAGG	1740
Qy	1741	GGTGCAAGAATAAATGCCATTACAAATTTCCACGATACACTCGCGCCAGTCAGACGATCTCA	1800
Db	1741	GGTGCAAGAATAAATGCCATTACAAATTTCCACGATACACTCGCGCCAGTCAGACGATCTCA	1800
Qy	1801	TTACACCACCATCATGACTACCATCATATTTCCATCATFACACCACCACCAAAACCCACATC	1860
Db	1801	TTACACCACCATCATGACTACCATCATATTTCCATCATFACACCACCACCAAAACCCACATC	1860
Qy	1861	CTACAGTTCACAGCCAGCGGTACTCTCGGGAGGAGCTGAAAGATGCGCGGCTGCGCCACTT	1920
Db	1861	CTACAGTTCACAGCCAGCGGTACTCTCGGGAGGAGCTGAAAGATGCGCGGCTGCGCCACTT	1920
Qy	1921	TGCGCTGGAGTGTGATTAATGGGTGATCGGCTGCACAAATTTACAGCATGCGCTTAGCAATTG	1980
Db	1921	TGCGCTGGAGTGTGATTAATGGGTGATCGGCTGCACAAATTTACAGCATGCGCTTAGCAATTG	1980
Qy	1981	GTGCTGCTTTTACTGAAGCTTTATCAAGTGGTTTAAGTACTCTGTGTGTGTGTCTGTCTGTC	2040
Db	1981	GTGCTGCTTTTACTGAAGCTTTATCAAGTGGTTTAAGTACTCTGTGTGTGTGTCTGTCTGTC	2040
Qy	2041	ATGAGTTGCGCTCATGAATFAGGTGACTTTGCTGTCTTCTACTAAAGCTGGCATCAGCGTTA	2100
Db	2041	ATGAGTTGCGCTCATGAATFAGGTGACTTTGCTGTCTTCTACTAAAGCTGGCATCAGCGTTA	2100
Qy	2101	ACGAGCTGTCTTTTATAATGCAATGTGACCATGTGCGGTATCTTGGAAATGCCAACAG	2160
Db	2101	ACGAGCTGTCTTTTATAATGCAATGTGACCATGTGCGGTATCTTGGAAATGCCAACAG	2160
Qy	2161	GAATTTTCATGATGTTGCTGTGTTGATATGGTACCTGAAATGCTGCACAAATGATGCTA	2220
Db	2161	GAATTTTCATGATGTTGCTGTGTTGATATGGTACCTGAAATGCTGCACAAATGATGCTA	2220
Qy	2221	GCTTATTTCATGATGTTGCTGTGTTGATATGGTACCTGAAATGCTGCACAAATGATGCTA	2280
Db	2221	GCTTATTTCATGATGTTGCTGTGTTGATATGGTACCTGAAATGCTGCACAAATGATGCTA	2280
Qy	2281	GTGACCATGGATGTAGCCGTGGGGTATTTCTTTTACAAATGCTGGGATGCTTTTGG	2340
Db	2281	GTGACCATGGATGTAGCCGTGGGGTATTTCTTTTACAAATGCTGGGATGCTTTTGG	2340
Qy	2341	GTTTGGAAATATGTTTACTTATTTCCATATTTGCAATATAATCGGTTTCTGATAAT	2400
Db	2341	GTTTGGAAATATGTTTACTTATTTCCATATTTGCAATATAATCGGTTTCTGATAAT	2400
Qy	2401	TCTAGTTAGGTTTAAATGCTAGATAGCTTAAAGCTGTCTATAGTTTCACTAGTGTAT	2460
Db	2401	TCTAGTTAGGTTTAAATGCTAGATAGCTTAAAGCTGTCTATAGTTTCACTAGTGTAT	2460
Qy	2461	AGGAGATGAGTTTGTATGCTGTACTATGACGCGTTTAAAGTTAGTGGGTTTGTGATTT	2520
Db	2461	AGGAGATGAGTTTGTATGCTGTACTATGACGCGTTTAAAGTTAGTGGGTTTGTGATTT	2520
Qy	2521	TTGTTATGAATATGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2580
Db	2521	TTGTTATGAATATGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2580
Qy	2581	TATTCATCTCTGGAGATATAATCTGTATGTGCAATTCACCGGTATACCGTTTATATG	2640
Db	2581	TATTCATCTCTGGAGATATAATCTGTATGTGCAATTCACCGGTATACCGTTTATATG	2640
Qy	2641	TAAACAAGATTTGGCATGACATGTCTGTATGTTTCAGGAAAAATGCTTTAATGCT	2700
Db	2641	TAAACAAGATTTGGCATGACATGTCTGTATGTTTCAGGAAAAATGCTTTAATGCT	2700
Qy	2701	TTTTTCAAGAACTAACACAGTTATTCCTATCTGGAATTTAGTCTCTGAAAGACTGCTG	2760
Db	2701	TTTTTCAAGAACTAACACAGTTATTCCTATCTGGAATTTAGTCTCTGAAAGACTGCTG	2760
Qy	2761	TGTTTGAAGTAATGAAGTGTGATGCAAGCTTAAATACCAAGAAAGCTTATCTGATTTA	2820
Db	2761	TGTTTGAAGTAATGAAGTGTGATGCAAGCTTAAATACCAAGAAAGCTTATCTGATTTA	2820

Db 2761 TGTTAGGAATAGAATTGTCATCGAACCCCTAAATAACCAAGAAAGCCTTATACTAGTAATTA 2820
QY 2821 ACACAAGAAATTAAGGGAGAAAAGAGAAATCTCAGAAATTTGGGAGGCCCATAGATCTCTAT 2880
Db 2821 AAAAAAAAAAAGAAAAAGAGAAATCTCAGAAATTTGGGAGGCCCATAGATCTCTAT 2880
QY 2881 AAAAATCACAAAATTTGTGTAAATTAGAGGGGAGAAAATTTAGAAATTAAGTGATATAAAGG 2940
Db 2881 AAAAAATCACAAAATTTGTGTAAATTAGAGGGGAGAAAATTTAGAAATTAAGTGATATAAAGG 2940
QY 2941 CAGAAATAGTATAGAGTACATTCATTAACATTTTTGTGCAGGATTTTCCCGTAAAAAC 3000
Db 2941 CAGAAATAGTATAGAGTACATTCATTAACATTTTTGTGCAGGATTTTCCCGTAAAAAC 3000
QY 3001 GTAGTGACACTCTCATATACTAATAATAGTGTACATTTAACTTTGTATATAACAGAAATCT 3060
Db 3001 GTFGTGACACTCTCATATACTAATAATAGTGTACATTTAACTTTGTATATAACAGAAATCT 3060
QY 3061 AAATATATTAAATGAATCAAGCAATATACACTTGCACCAGAAATTTGGAATTTCAAATG 3120
Db 3061 AAATATATTAAATGAATCAAGCAATATACACTTGCACCAGAAATTTGGAATTTCAAATG 3120
QY 3121 TTCTGCGGGTTATATACCAAGTACAGTACAGTACAGTACAGTATGATGATCACCAGACTGGGT 3180
Db 3121 TTCTGCGGGTTATATACCAAGTACAGTACAGTACAGTACAGTATGATGATCACCAGACTGGGT 3180
QY 3181 TATTGCCAAGTTATATATACCAAAAGCTGTATGACTGCGATGTTCTGCTTACCTGGTTTGA 3240
Db 3181 TATTGCCAAGTTATATATACCAAAAGCTGTATGACTGCGATGTTCTGCTTACCTGGTTTGA 3240
QY 3241 CAATAATTATCAGAGTAGTAACTTTGATATATAGAGGATATTAATACTACACTAAGTA 3300
Db 3241 CAATAATTATCAGAGTAGTAACTTTGATATATAGAGGATATTAATACTACACTAAGTA 3300
QY 3301 TCATTGTGATTCGATTACAGAAAGTACTTTGATATCTCAGTGCCTTCAGTGCATCATTTGT 3360
Db 3301 TCATTGTGATTCGATTACAGAAAGTACTTTGATATCTCAGTGCCTTCAGTGCATCATTTGT 3360
QY 3361 GAGCAATGCTTTATATACGGTACTGTAGCCATACAGCCCTGCTGCTGGCATCTCTA 3420
Db 3361 GAGCAATGCTTTATATACGGTACTGTAGCCATACAGCCCTGCTGCTGGCATCTCTA 3420
QY 3421 GATGTTCTTTTTCACAAATAAATTCCTTATATACAGTTG 3461
Db 3421 GATGTTCTTTTTCACAAATAAATTCCTTATATACAGTTG 3461

RESULT 2
LOCUS AX207205 3461 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0155178.
VERSION AX207205
KEYWORDS AX207205.1 GI:15394960
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 3461)
AUTHORS Goddard,A., Guney,A.L., Smith,V., Hong,J.A. and de Sauvage,F.
TITLE Compositions and methods for treatment of cancer
JOURNAL Patent: WO 0155178-A 1 02-AUG-2001;

FEATURES
source Location/Qualifiers
1..3461 /organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 1057 a 679 c 727 g 998 t
ORIGIN

Query Match 98.0%; Score 3391; DB 6; Length 3461;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3441; Conservative 0; Mismatches 0; Indels 20; Gaps 3

RESULT 2				PAT 30-AUG-2001
AX207205			linear	
AX207206	3461 bp	DNA		
locus	AX207205	patent W00155178.		

KEYWORDS
human.
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 3461)
Goddard, A., Gurney, A. L., Smith, V., Hongo, J. A. and de Sauvage, F. F.
Compositions and methods for treatment of cancer
FILE
11-02-AUG-2001.

JOURNAL
FACENS.
GENENTECH, INC. (US)
Location/Qualifiers
1. .3461
source

```

/db_xref="taxon:9606"
998 t

```

BASE COUNT	1057 a	012 c
ORIGIN		
OR 6:	Length 34617	

Query Match	98.0%;	Score 3391;	25 07
Best Local Similarity	99.4%;	Pred. No. 0;	
Matches 3441; Conservative	0;	Mismatches	0;
		Indels	20;
		Gaps	3;

QY	1	CTCGTGGCAATTCCGACAGACCCTGTTCGCGCTGTTAGAGATTTCTCGAAGACA	60
DB	1	CTCGTGGCAATTCCGACAGAGACGGGTGTTCGCGCTGGTAGATTTCTCGAAGACA	60
QY	61	CCAGTGGGCCGTGTGGAACAAACACTGCGCGCGTGCCGGCGGTGGACAAACGAGGC	120
DB	61	CCAAGTGGGCCGTGTGGAACAAACACTGCGCGCGTGCCGGCGGTGGACAAACGAGGC	120
QY	121	GCGGAGACGAGGCCAATGGCAGAGAAATTTATCTGTAATCTTGATCCTGACCTTTGCC	180
DB	121	GCGGAGACGAGGCCAATGGCAGAGAAATTTATCTGTAATCTTGATCCTGACCTTTGCC	180
QY	181	TCTCTGTACAAAATCCCCTTCATGAACTAAAGCAGCTGCTTTCCGCCACACACTGAGA	240
DB	181	TCTCTGTACAAAATCCCCTTCATGAACTAAAGCAGCTGCTTTCCGCCACACACTGAGA	240
QY	241	AAATAGTCCGAATGGGAATCTGGCATTAATGTTGACTTGGCAATTTCCACACGGCAAT	300
DB	241	AAATAGTCCGAATGGGAATCTGGCATTAATGTTGACTTGGCAATTTCCACACGGCAAT	300
QY	301	ATCATCTACACAGCTTTTCTACGCTATGGAGAAATAATTTCTTTGTTCAGTTGAAGGT	360
DB	301	ATCATCTACACAGCTTTTCTACGCTATGGAGAAATAATTTCTTTGTTCAGTTGAAGGT	360
QY	361	TCAGAAAATTTACTTCAAATAATAGGCATAGATAAGATTAAAGAAATTCATACACCATG	420
DB	361	TCAGAAAATTTACTTCAAATAATAGGCATAGATAAGATTAAAGAAATTCATACACCATG	420
QY	421	ACCACGACCATCACTCAGACCACGAGCATCACTCAGACCATGAGCGTCACTCAGACCATG	480
DB	421	ACCACGACCATCACTCAGACCACGAGCATCACTCAGACCATGAGCGTCACTCAGACCATG	480
QY	481	AGCATCACTCAGACCACGAGCATCACTCTGACCATGATCATCTCACCAATAATCATG	540
DB	481	AGCATCACTCAGACCACGAGCATCACTCTGACCATGATCATCTCACCAATAATCATG	540
QY	541	CTCGTTCTGTTAAAAATAGCCANAAGCTTTTGGCCCCAGACCATGACTCAGATGATTCAG	600
DB	523	CTCGTTCTGTTAAAAATAGCCANAAGCTTTTGGCCCCAGACCATGACTCAGATGATTCAG	582
QY	601	CFAAGATCTCTAGAACAGCCAGGGGAAAGAGCTCACCGACCAAGACATGCCAGTGGTA	660
DB	583	GTAAGATCCTAGAACAGCCAGGGGAAAGAGCTCACCGACCAAGACATGCCAGTGGTA	642
QY	661	GAAGGAATGTCNAGGACAGTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT	720
DB	643	GAAGGAATGTCNAGGACAGTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT	702
QY	721	TCCTCTGAAGGAACCTCACTTTCTAGAGACAATAGAGACTCCAAGACTTGAAGACTTTC	780
DB	703	TCCTCTGAAGGAACCTCACTTTCTAGAGACAATAGAGACTCCAAGACTTGAAGACTTTC	762
QY	781	CCAAAGATGTAGGAGCTCCACTCCACCAGTGTACATCAAGAGCGGGGTGACCGGC	840
DB	763	CCAAAGATGTAGGAGCTCCACTCCACCAGTGTACATCAAGAGCGGGGTGACCGGC	822
QY	841	TGCGTGGTAGGAAAAACAAATGAATCTGTGAGTGAGCCCGGAAAGGCTTTATGTAATCCA	900
DB	823	TGCGTGGTAGGAAAAACAAATGAATCTGTGAGTGAGCCCGGAAAGGCTTTATGTAATCCA	882
QY	901	GAAACACAAATGAAATCTCTAGGAGTGTCTCAATGTCATCAAGACTTACTGACATCTCATG	960
DB	883	GAAACACAAATGAAATCTCTAGGAGTGTCTCAATGTCATCAAGACTTACTGACATCTCATG	942
QY	961	GCATGGGCATCCAGGTTCCGCTGAATGCAACAGAGTTCAACTATCTCTGTCCAGCCATCA	1020
DB	943	GCATGGGCATCCAGGTTCCGCTGAATGCAACAGAGTTCAACTATCTCTGTCCAGCCATCA	1002
QY	1021	TCACCAAAATGTAGCTAGATCTTGTCTGATTCATACAGTGAAAGAGGCTGAAATCC	1080
DB	1003	TCACCAAAATGTAGCTAGATCTTGTCTGATTCATACAGTGAAAGAGGCTGAAATCC	1062

QY	1081	CTCAAAGACCTATTTCATTAACAATAGCGTGGGTTGGTGTTTTATAGCCATTTCCATCA	1140
DB	1063	CTCAAAGACCTAATTCATTACAAATAGCGTGGGTTGGTGTTTTATAGCCATTTCCATCA	1122
QY	1141	TCAGTTTTCCGTCTCTGCTGGGGGTTATCTTAGTGCCCTCATGAATCGGGTGTTTTTCA	1200
DB	1123	TCAGTTTTCCGTCTCTGCTGGGGGTTATCTTAGTGCCCTCATGAATCGGGTGTTTTTCA	1182
QY	1201	AATTTCTCCTGAGTTCTCTGTGGCACTGGCCGTTGGGACTTTGAGTGGTGATGCTTTTT	1260
DB	1183	AATTTCTCCTGAGTTCTCTGTGGCACTGGCCGTTGGGACTTTGAGTGGTGATGCTTTTT	1242
QY	1261	TACACCTTTCTCCACATTCATCTGCAAGTCACCACCATAGTCATAGGCATCAAGAACCAG	1320
DB	1243	TACACCTTTCTCCACATTCATCTGCAAGTCACCACCATAGTCATAGGCATCAAGAACCAG	1302
QY	1321	CAATGGAATGAAAAGAGGACCACTTTTCAGTCTCTGCTTCTCAAAACATAAGAAGAA	1380
DB	1303	CAATGGAATGAAAAGAGGACCACTTTTCAGTCTCTGCTTCTCAAAACATAAGAAGAA	1362
QY	1381	GTGCCATTTTTCATTCACAGTGGAAAGGCTCTAACAGCTCTAGGAGGCTGTATTTCAATG	1440
DB	1363	GTGCCATTTTTCATTCACAGTGGAAAGGCTCTAACAGCTCTAGGAGGCTGTATTTCAATG	1422
QY	1441	TTCTCTGTGAACATGTCCTCACATGTATGATCAACAANTTTAAGATAGAAGAAAAAGATC	1500
DB	1423	TTCTCTGTGAACATGTCCTCACATGTATGATCAACAANTTTAAGATAGAAGAAAAAGATC	1482
QY	1501	AGAGAANCTGGAATGATGATGATGATGGAGATTAAGAAGCAGTGTGCCAGTATGAAT	1560
DB	1483	AGAGAANCTGGAATGATGATGATGATGGAGATTAAGAAGCAGTGTGCCAGTATGAAT	1542
QY	1561	CTCAACTTTCAACAATGAGGAGAAAGTAGATACAGATGATCGAAGCTGAAGGCTATTTAC	1620
DB	1543	CTCAACTTTCAACAATGAGGAGAAAGTAGATACAGATGATCGAAGCTGAAGGCTATTTAC	1602
QY	1621	GAGCAGACTCAAGAAGCCCCCTCCACTTTGATTTCTCAGCAGCCTGCGAGTCTCGGAAGAAG	1680
DB	1603	GAGCAGACTCAAGAAGCCCCCTCCACTTTGATTTCTCAGCAGCCTGCGAGTCTCGGAAGAAG	1662
QY	1681	NAGAGGTATGNTAGCTCATGCTCATCCACAGGAAGTCTACAAATGAATATGTACCCNAG	1740
DB	1663	NAGAGGTATGNTAGCTCATGCTCATCCACAGGAAGTCTACAAATGAATATGTACCCNAG	1722
QY	1741	GGTGCAGAANAATAATGCCATTTCCAGATTTCCACCATACACTCGGCCAGTCAAGCATCTCA	1800
DB	1723	GGTGCAGAANAATAATGCCATTTCCAGATTTCCACCATACACTCGGCCAGTCAAGCATCTCA	1782
QY	1801	TTCCACCACCATCATGACTACCATCATATTTCCCATCATCACACCACCAAAACACCATC	1860
DB	1783	TTCCACCACCATCATGACTACCATCATATTTCCCATCATCACACCACCAAAACACCATC	1842
QY	1861	CTCAGACTCAGCCAGCGCTACTCTCGGGAGGAGCTGAAAGATGCGCGGCTGCGCACTT	1920
DB	1843	CTCAGACTCAGCCAGCGCTACTCTCGGGAGGAGCTGAAAGATGCGCGGCTGCGCACTT	1902
QY	1921	TGGCCTGGATGGTGATAATGGGTGATGGCGTGCACATTTTCAGCGATGGCCTAGCAATTTG	1980
DB	1903	TGGCCTGGATGGTGATAATGGGTGATGGCGTGCACATTTTCAGCGATGGCCTAGCAATTTG	1962
QY	1981	GTGCTGCTTTTACTGAAGGCTTATCAAGTGGTTTTAAGTACTTCTGTGTGCTGCTCTGTC	2040
DB	1963	GTGCTGCTTTTACTGAAGGCTTATCAAGTGGTTTTAAGTACTTCTGTGTGCTGCTCTGTC	2022
QY	2041	ATGAGTTGGCTTCATGAATTAGGTGACTTTTGCTTCTACTAAAGGCTGGCATGACCGTTA	2100
DB	2023	ATGAGTTGGCTTCATGAATTAGGTGACTTTTGCTTCTACTAAAGGCTGGCATGACCGTTA	2082
QY	2101	AGCAGGCTGCTCTTTATAAATGCAATTTGTACGCCATGCTGGCGTATCTTGGAAATGCAACAG	2160
DB	2083	AGCAGGCTGCTCTTTATAAATGCAATTTGTACGCCATGCTGGCGTATCTTGGAAATGCAACAG	2142
QY	2161	GAATTTPTCATTTGGTCAATTAACCTCAAAAATCTTTCTATGTGGATATATTGCACTTACTGCTG	2220

Db 3321 CAAATATTACAGAGTAGTAAACCTTGGATATCTCTCAGTCTCTCAGTCTATCATTTGT 3360
 QY 3301 TCATTTGATTGCATTCAGAAAGTACTTTTGATATCTCTCAGTCTCTCAGTCTATCATTTGT 3340
 Db 3281 TCATTTGATTGCATTCAGAAAGTACTTTTGATATCTCTCAGTCTCTCAGTCTATCATTTGT 3340
 QY 3361 GAGCAATGCTTTTATATACGCTACTGTAGCCACTACTAGGCTGTCTGGCAATCTCTA 3420
 Db 3341 GAGCAATGCTTTTATATACGCTACTGTAGCCACTACTAGGCTGTCTGGCAATCTCTA 3400
 QY 3421 GATGTTCTTTTACACATAAATTCCTTTATATCAGCTTG 3461
 Db 3401 GATGTTCTTTTACACATAAATTCCTTTATATCAGCTTG 3441

RESULT 3
 LOCUS AX207207 2776 bp DNA linear PAT 30-AUG-2001
 DEFINITION Sequence 3 from Patent WO0155178.
 ACCESSION AX207207
 VERSION AX207207.1 GI:15394961
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 2776)
 Goddard, A., Gurney, A.L., Smith, V., Hongo, J.A. and de Sauvage, F.
 Compositions and methods for treatment of cancer
 Patent: WO 0155178-A 3 02-AUG-2001.
 GENENTECH, INC. (US)

FEATURES
 Location/Qualifiers
 1..2776
 /organism="Homo sapiens"
 /db_xref="taxon:9606" 778 t
 823 a 583 c 592 g

BASE COUNT 823 a 583 c 592 g
 ORIGIN

Query Match 79.8%; Score 2761.2; DB 6; Length 2776;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2766; Conservative 0; Mismatches 8; Indels 0; Gaps

QY 68 GCCCGTGTGGAAACCAACCTGCGCGTGGCGGGCGCGTGGACACGAGCGCGGAGA 127
 Db 3 GCCCGTGTGGAAACCAACCTGCGCGTGGCGGGCGCGTGGACACGAGCGCGGAGA 62
 QY 128 CGAAGCGCAATGCGGAGGAAGTTATCTTAATCTTGATCTGACCTTTGCCCTCTCTGT 187
 Db 63 CGAAGCGCAATGCGGAGGAAGTTATCTTAATCTTGATCTGACCTTTGCCCTCTCTGT 122
 QY 188 CACAAATCCCTTCATGAACCTAAAGCAGCTGCTTCCCGACACCACTCAGAAAAATTAG 247
 Db 123 CACAAATCCCTTCATGAACCTAAAGCAGCTGCTTCCCGACACCACTCAGAAAAATTAG 382
 QY 248 TCCGAATGGGAATCTGGCAATTAATGTTGACTTGGCAATTTCCACCGGCAATATCATCT 307
 Db 183 TCCGAATGGGAATCTGGCAATTAATGTTGACTTGGCAATTTCCACCGGCAATATCATCT 242
 QY 308 ACAACAGCTTTTCTACCGTATGGAGAAAATAATCTTTCTCAGTTGAAGGTTTCAGAAA 367
 Db 243 ACAACAGCTTTTCTACCGTATGGAGAAAATAATCTTTCTCAGTTGAAGGTTTCAGAAA 302
 QY 368 ATTACTTCAAAATATAGGCATAGATGAAGATTAAAGAAATCCATATACACCATGACACGA 427
 Db 303 ATTACTTCAAAATATAGGCATAGATGAAGATTAAAGAAATCCATATACACCATGACACGA 362
 QY 428 CCATCTACAGACACGAGCATCACTCAGACCATGAGCGTCACTCAGACCATGAGCATCA 487
 Db 363 CCATCTACAGACACGAGCATCACTCAGACCATGAGCGTCACTCAGACCATGAGCATCA 422
 QY 488 CTCAGACCGAGCATCACTCTGACCATGATCATCACTCTCAGCATATCATCTGCTGCTTC 547
 Db 423 CTCAGACCGAGCATCACTCTGACCATGATCATCACTCTCAGCATATCATCTGCTGCTTC 482

|||||
Db 2643 GAACTAACACACTTATTCCTTACTGATTTTAGTCTCTGAAGAACTGCTGGTGTAG 2702
Qy 2768 GAATAGATGTCATGAAGCTTAAATACCAAGAAAGCTTATCTACTGAATTTAAAGCAAAG 2827
Db 2703 GAATAGATGTCATGAAGCTTAAATACCAAGAAAGCTTATCTACTGAATTTAAAGCAA 2762
Qy 2828 AAATAAGCAGAA 2841
Db 2763 AAAAAAAAAAAAA 2776

RESULT 4
HSU41060
LOCUS
DEFINITION Homo sapiens estrogen regulated LIV-1 protein (LIV-1) mRNA, PRI 08-FEB-2001
complete cds.
ACCESSION U41060
VERSION U41060.2 GI:12711792
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2744)
Green, C. and Morgan, H.
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (08-FEB-2001) Biochemistry, University of Liverpool, P.O.
Box 147, Liverpool L69 3BX, UK
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (21-NOV-1995) Biochemistry, University of Liverpool, P.O.
Box 147, Liverpool L69 3BX, UK
COMMENT On Feb 8, 2001 this sequence version replaced gi:1256000.
FEATURES
source
1..2744
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="MCF-7 human breast cancer cell line"
/note="estrogen induced mRNA"
1..2744
/gene="LIV-1"
138..2387
/gene="LIV-1"
/note="breast cancer"
/codon_start=1
/product="estrogen regulated LIV-1 protein"
/protein_id="AAA96258.2"
/db_xref="GI:12711793"
translation="MARLKVILITFALSVTNPLHKLAAAFPTOTKISPNWESGI
NVDLAISTROYHLOOLFVRYGENSLVGEFRKLQNGIDKIKRIHHDHSDH
EHSHERSHDHSDEHSDHHAASGKNRKALCPDHDSDSGKDPNRSQCKGA
HRPEHASGRNRKDVSASEVTSTVNTVSEGTHTLETETPRPKLPKDVSSSTPP
SVTSKSRVLRAGRTNESVSPRGMYSRNTNENPQCFNASKLLTSHGMGIQVPL
NATEFNLYCPAIIHQIDARSLIHTSEKRAEIPPKTYLQIAWVGFTAISITSL
LGVLYPLMNRVFKLLSLVALAVGTLSGDAFLHLLPHSHASHHSHEEPANEM
KCPLESHLSSONIESAYFDSWTGKLTALGLYFMFLVHVLTLIKQFKKKKQK
KPNDDVEIKQLSKYESQLSTNEEKVDTRTEGYLRADSOEFSHQOQPAVLEE
EYVIAHAHPQYNEYVPRCKNCKSHFHDTLGSDDLIHHHDYHHLHHHHQK
HHPHSQSRYSRELDAGVATLWVIMGDLNFSGLAIGAAFTGSGLSLTSV
AVFCHPELHGLDFVLLKAGMTVKQVLYNALSLAMLAYLGMATGFIGHYAENYSW
IFALTAGLEMYVALVDWVPEMLHNDASDHGCSRWGYFFLQNGMLLIGFIMLLISIE
HKIVFRNF"
BASE COUNT 786 a 585 c 601 g 772 t
ORIGIN
Query Match 78.5%; Score 2716; DB 9; Length 2744;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2744; Conservative 0; Mismatches 0; Indels 18; Gaps 1;
Qy 1 CTCGTGCCGAATTCGGCAGACAGCGGTGTTCCGCCCTGGTAGAGATTCTCGAAGACA 60
|||||

Db 1 CTCGTGCCGAATTCGGCAGACAGCGGTGTTCCGCCCTGGTAGAGATTCTCGAAGACA 60
Qy 61 CCAGTGGGCCGCGTGTGGAAACCAACCTGCGCGGTGGCGGGCGGTGGGACACAGAGGCC 120
Db 61 CCAGTGGGCCGCGTGTGGAAACCAACCTGCGCGGTGGCGGGCGGTGGGACACAGAGGCC 120
Qy 121 GCGGAGACGAAGCGGCAATGCGGAGGAAGTTATCTGTAAATCTTTGATCCTGACCTTTGCC 180
Db 121 GCGGAGACGAAGCGGCAATGCGGAGGAAGTTATCTGTAAATCTTTGATCCTGACCTTTGCC 180
Qy 181 TCTCTGTACAAAATCCCTTCATGAACATAAAGCAGCTGCTTTCCGCCAGACCACTGAGA 240
Db 181 TCTCTGTACAAAATCCCTTCATGAACATAAAGCAGCTGCTTTCCGCCAGACCACTGAGA 240
Qy 241 AATTAGTCGAATTTGGGAATCTGGCAATTAATGTTGACTTGGCAATTTCCACACGCGCAAT 300
Db 241 AATTAGTCGAATTTGGGAATCTGGCAATTAATGTTGACTTGGCAATTTCCACACGCGCAAT 300
Qy 301 ATCATCTACAACAGCTTTTCTACCGCTATGGAGAAAATAATTTCTTGTCTAGTTGAAGGGT 360
Db 301 ATCATCTACAACAGCTTTTCTACCGCTATGGAGAAAATAATTTCTTGTCTAGTTGAAGGGT 360
Qy 361 TCAGAAAATTAATCTCAAAAATATAGGCATAGATAAGATTAAAGAAATCCATATACACCATG 420
Db 361 TCAGAAAATTAATCTCAAAAATATAGGCATAGATAAGATTAAAGAAATCCATATACACCATG 420
Qy 421 ACCAGGACCATCTCAGACCACGAGCATCACTCAGACCACGAGCGCTCACTCAGACCATG 480
Db 421 ACCAGGACCATCTCAGACCACGAGCATCACTCAGACCACGAGCGCTCACTCAGACCATG 480
Qy 481 AGCATCACTCAGACCACGAGCATCACTCTGACCATGATCATCACTCTCACCATAATCATG 540
Db 481 AGCATCACTCAGACCACGAGCATCACTCTGACCATGATCATCACTCTCACCATAATCATG 540
Qy 541 CTGCTTCTGGTAAATTAAGGAAAAGCTCTTTGCCAGACCATGACTCAGATGATTCAG 600
Db 541 CTGCTTCTGGTAAATTAAGGAAAAGCTCTTTGCCAGACCATGACTCAGATGATTCAG 600
Qy 583 GTAAAGATCTCTAGAAACAGCCAGGGAAGGAGCTCACCCAGACCAATGCCAGTGGTA 642
Db 583 GTAAAGATCTCTAGAAACAGCCAGGGAAGGAGCTCACCCAGACCAATGCCAGTGGTA 642
Qy 661 GAAGGAATGTCAAGGACAGTGTAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 720
Db 661 GAAGGAATGTCAAGGACAGTGTAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 720
Qy 702 GAAGGAATGTCAAGGACAGTGTAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 702
Db 702 GAAGGAATGTCAAGGACAGTGTAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 702
Qy 721 TCTCTGAAGGAATCTCACTTTCTAGAGACAATAGAGACTCCAGAGCTCGAAAACCTCTTCC 780
Db 721 TCTCTGAAGGAATCTCACTTTCTAGAGACAATAGAGACTCCAGAGCTCGAAAACCTCTTCC 780
Qy 703 TCTCTGAAGGAATCTCACTTTCTAGAGACAATAGAGACTCCAGAGCTCGAAAACCTCTTCC 762
Db 703 TCTCTGAAGGAATCTCACTTTCTAGAGACAATAGAGACTCCAGAGCTCGAAAACCTCTTCC 762
Qy 781 CCAAGATGTAAAGCAGCTCCACTCCACCAGTGTCCACATCAAAGAGCGGGGTGAGCGGCG 840
Db 781 CCAAGATGTAAAGCAGCTCCACTCCACCAGTGTCCACATCAAAGAGCGGGGTGAGCGGCG 840
Qy 841 TGGCTGTAGGAAAACAAATGAATCTGTGAGTGAGCGGCGGAAAAGGCTTTATGTATTCCA 900
Db 841 TGGCTGTAGGAAAACAAATGAATCTGTGAGTGAGCGGCGGAAAAGGCTTTATGTATTCCA 900
Qy 882 TGGCTGTAGGAAAACAAATGAATCTGTGAGTGAGCGGCGGAAAAGGCTTTATGTATTCCA 882
Db 882 TGGCTGTAGGAAAACAAATGAATCTGTGAGTGAGCGGCGGAAAAGGCTTTATGTATTCCA 882
Qy 901 GAAACAAAATGAAAATCTCAGGAGTGTTCATATGATCAATGATCAATGATCAATGATCAATG 960
Db 901 GAAACAAAATGAAAATCTCAGGAGTGTTCATATGATCAATGATCAATGATCAATGATCAATG 960
Qy 883 GAAACAAAATGAAAATCTCAGGAGTGTTCATATGATCAATGATCAATGATCAATGATCAATG 942
Db 883 GAAACAAAATGAAAATCTCAGGAGTGTTCATATGATCAATGATCAATGATCAATGATCAATG 942
Qy 961 GCATGGGCATCCAGGTTCCGCTGAATGCAACAGAGTTCAACTATCTCTCTCCAGGCATCA 1020
Db 961 GCATGGGCATCCAGGTTCCGCTGAATGCAACAGAGTTCAACTATCTCTCTCCAGGCATCA 1020
Qy 943 GCATGGGCATCCAGGTTCCGCTGAATGCAACAGAGTTCAACTATCTCTCTCCAGGCATCA 1002
Db 943 GCATGGGCATCCAGGTTCCGCTGAATGCAACAGAGTTCAACTATCTCTCTCCAGGCATCA 1002
Qy 1021 TCAACCAATTTGATGCTAGATCTTCTCATCAATCAAGTGAAGAGGCTGAATCC 1080
Db 1021 TCAACCAATTTGATGCTAGATCTTCTCATCAATCAAGTGAAGAGGCTGAATCC 1080
Qy 1003 TCAACCAATTTGATGCTAGATCTTCTCATCAATCAAGTGAAGAGGCTGAATCC 1062
Db 1003 TCAACCAATTTGATGCTAGATCTTCTCATCAATCAAGTGAAGAGGCTGAATCC 1062
Qy 1081 CTCGAGGACCTATTCTTCAATACAAATAGCCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1140
Db 1081 CTCGAGGACCTATTCTTCAATACAAATAGCCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1140
Qy 1063 CTCGAGGACCTATTCTTCAATACAAATAGCCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1122
Db 1063 CTCGAGGACCTATTCTTCAATACAAATAGCCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1122

QY 1141 TCAGTTTCCCTGCTCTGCTGGGGTATCTTAGTGGCTCTCATGAATCGGGTGTCTTTTCA 1200
DB 1123 TCAGTTTCCCTGCTCTGCTGGGGTATCTTAGTGGCTCTCATGAATCGGGTGTCTTTTCA 1182
QY 1201 AATTTTCCTGAGTTTCTGCTGGCACTGGCCGTTGGGACTTTGAGTGGTGATGCTTTT 1260
DB 1183 AATTTTCCTGAGTTTCTGCTGGCACTGGCCGTTGGGACTTTGAGTGGTGATGCTTTT 1242
QY 1261 TACACCTTCTTCCACATTTCTATGCAAGTCACCACTAGTCATAGCCATGAAGAACCCAG 1320
DB 1243 TACACCTTCTTCCACATTTCTATGCAAGTCACCACTAGTCATAGCCATGAAGAACCCAG 1302
QY 1321 CAATGGAATGAAAGAGGACCACTTTTCACTCATCTGCTCTCAAAACATAGAAGAA 1380
DB 1303 CAATGGAATGAAAGAGGACCACTTTTCACTCATCTGCTCTCAAAACATAGAAGAA 1362
QY 1381 GTGCCTATTTTGAATTCACACGTGGAAGGGTCTAACAGCTCTAGGAGCCCTGTATTTCACTG 1440
DB 1363 GTGCCTATTTTGAATTCACACGTGGAAGGGTCTAACAGCTCTAGGAGCCCTGTATTTCACTG 1422
QY 1441 TTCTTGTGGAACATGCTCTCACATTTGATCAACAAATTAAGATAGAAGAAAAGATC 1500
DB 1423 TTCTTGTGGAACATGCTCTCACATTTGATCAACAAATTAAGATAGAAGAAAAGATC 1482
QY 1501 ACAAGAACCTGAAATGATGATGCTGGAGATTAAGAAGCAGTTGTCCAACTATGAAT 1560
DB 1483 AGAAGAACCTGAAATGATGATGCTGGAGATTAAGAAGCAGTTGTCCAACTATGAAT 1542
QY 1561 CTCAACTTTCAACAAATGAGGAGAACTAGATACAGATGATGCAACTGAAGGCTATTTAC 1620
DB 1543 CTCAACTTTCAACAAATGAGGAGAACTAGATACAGATGATGCAACTGAAGGCTATTTAC 1602
QY 1621 GAGCAGATCACAGAGCCCTCCACTTTGATTTCTCAGCAGCTTGCAAGTCTTGGAAGAG 1680
DB 1603 GAGCAGATCACAGAGCCCTCCACTTTGATTTCTCAGCAGCTTGCAAGTCTTGGAAGAG 1662
QY 1681 AAGAGTCTATGATAGCTCATGCTCATCCAGGAGCTTACAATGAATATGTACCCAGAG 1740
DB 1663 AAGAGTCTATGATAGCTCATGCTCATCCAGGAGCTTACAATGAATATGTACCCAGAG 1722
QY 1741 GGTGCAAGATTAATGGCATTCATTTCCACGATACACTCGCCGCTCAGACGATCTCA 1800
DB 1723 GGTGCAAGATTAATGGCATTCATTTCCACGATACACTCGCCGCTCAGACGATCTCA 1782
QY 1801 TTCACACCATCATGACTTACCATCATATTCCTCATCATCACACCAACCAACCCATC 1860
DB 1783 TTCACACCATCATGACTTACCATCATATTCCTCATCATCACACCAACCAACCCATC 1842
QY 1861 CTCACAGTCACAGCAGCGCTACTCTCGGGAGAGCTGAAAGATCGCCGGCTGCCACTT 1920
DB 1843 CTCACAGTCACAGCAGCGCTACTCTCGGGAGAGCTGAAAGATCGCCGGCTGCCACTT 1902
QY 1921 TGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980
DB 1903 TGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1962
QY 1981 GTGCTGCTTTTACTGAAGGCTTATCAAGTGGTTTAAAGTACTTCTGTGCTGTGCTGTC 2040
DB 1963 GTGCTGCTTTTACTGAAGGCTTATCAAGTGGTTTAAAGTACTTCTGTGCTGTGCTGTC 2022
QY 2041 ATGAGTTGCCATGAATAGTGAATAGTGAATAGTGAATAGTGAATAGTGAATAGTGAAT 2100
DB 2023 ATGAGTTGCCATGAATAGTGAATAGTGAATAGTGAATAGTGAATAGTGAATAGTGAAT 2082
QY 2101 AGCAGCTGCTCTTTTATATATGATGTCAGCAGCTGTCGCGTATCTTGGAAATGCAACAG 2160
DB 2083 AGCAGCTGCTCTTTTATATATGATGTCAGCAGCTGTCGCGTATCTTGGAAATGCAACAG 2142
QY 2161 GAAATTTTCATTTGCTATTTGCTGAAATATTTTCTATGATGATATTTGCACTTACTGCTG 2220
DB 2143 GAAATTTTCATTTGCTATTTGCTGAAATATTTTCTATGATGATATTTGCACTTACTGCTG 2202

QY 2221 GCTTATTCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTA 2280
DB 2203 GCTTATTCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTA 2262
QY 2281 GTGACCATGGATGTAGCGCTGGGGTATTTCTTTTACAGAATGCTGGGATGCTTTTGG 2340
DB 2263 GTGACCATGGATGTAGCGCTGGGGTATTTCTTTTACAGAATGCTGGGATGCTTTTGG 2322
QY 2341 GTTTTGGAAATATGTTACTTATTTCCATATTTGAACATATAAATCGTGTTCGTATAAAT 2400
DB 2323 GTTTTGGAAATATGTTACTTATTTCCATATTTGAACATATAAATCGTGTTCGTATAAAT 2382
QY 2401 TCTAGTTAAGGTTTAAATGCTAGAGTACGTTAAAAAGTTGTCATAGTTTTCAGTAGGTCAT 2460
DB 2383 TCTAGTTAAGGTTTAAATGCTAGAGTACGTTAAAAAGTTGTCATAGTTTTCAGTAGGTCAT 2442
QY 2461 AGGAGATGAGTTTGTATGCTGCTACTATGCGAGCGTTTAAAGTTAGTGGGTTTGTGATTT 2520
DB 2443 AGGAGATGAGTTTGTATGCTGCTACTATGCGAGCGTTTAAAGTTAGTGGGTTTGTGATTT 2502
QY 2521 TTGTATTGAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
DB 2503 TTGTATTGAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2562
QY 2581 TATTCATCTTGGAGATAAAATCTGTATGCAATTCACCGCTATTACCAGTTTATTATG 2640
DB 2563 TATTCATCTTGGAGATAAAATCTGTATGCAATTCACCGCTATTACCAGTTTATTATG 2622
QY 2641 TAAACAGAGATTTGGCATGACATGTTCTGTATGTTTCAGGGAAAAATGCTTTTAATGCT 2700
DB 2623 TAAACAGAGATTTGGCATGACATGTTCTGTATGTTTCAGGGAAAAATGCTTTTAATGCT 2682
QY 2701 TTTTCAAGAACTAAACACAGTTTATCTATCTAGTGGATTTTAGCTCTCTGAAGAACTGCTGG 2760
DB 2683 TTTTCAAGAACTAAACACAGTTTATCTATCTAGTGGATTTTAGCTCTCTGAAGAACTGCTGG 2742
QY 2761 TG 2762
DB 2743 TG 2744

RESULT 5
176892
LOCUS I76892 2404 bp DNA linear PAT 03-APR-1998
DEFINITION Sequence 3 from patent US 5693465.
ACCESSION I76892
VERSION I76892.1 GI:3013046
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2404)
AUTHORS Manning D.Lockwood., Nicholson R.Ian., Gee,J.Margaret.Wendy. and Green,C.Douglas.
TITLE Methods for predicting the behaviour of breast tumours
JOURNAL Patent: US 5693465-A 3 02-DEC-1997;
FEATURES
source Location/Qualifiers
i. .2404
/organism="unknown"
BASE COUNT 726 a 411 c 492 g 775 t
ORIGIN

Query Match 65.0%; Score 2249; DB 6; Length 2404;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 2341; Conservative 0; Mismatches 0; Indels 62; Gaps 3;

QY 1119 GGTTTTATAGCCATTTCCATCATCAGTTCCCTGCTCTGCTGGGGTATCTTAGTGCCT 1178
DB 1 GGTTTTATAGCCATTTCCATCATCAGTTCCCTGCTCTGCTGGGGTATCTTAGTGCCT 60
QY 1179 CTCATGAATCGGGTGTCTTTTCAAAATTTCCCTGAGTTTCTGTCGCACTGGCCGTTGG 1238
DB 61 CTCATGAATCGGGTGTCTTTTCAAAATTTCCCTGAGTTTCTGTCGCACTGGCCGTTGG 120

QY 1239 ACTTTGAGTGGTGATGCTTTTTTACACCTCTCTCCACATCTCTCATGCAAGTCACCCACAT 1298
DB 121 ACTTTGAGTGGTGATGCTTTTTTACACCTCTCTCCACATCTCTCATGCAAGTCACCCACAT 180
QY 1299 AGTCATAGCCATGAAGAACCCAGCAATGGAATGAAAGAGAGGACCCACTTTTCAGTCATCTG 1358
DB 181 AGTCATAGCCATGAAGAACCCAGCAATGGAATGAAAGAGAGGACCCACTTTTCAGTCATCTG 240
QY 1359 TCTTCTCAAAACATAGAGAAAGTGCCTATTTTCATTCCACGTCGGAAGGCTTAACAGCT 1418
DB 241 TCTTCTCAAAACATAGAGAAAGTGCCTATTTTCATTCCACGTCGGAAGGCTTAACAGCT 300
QY 1419 CTAGAGGCCCTGATTTTCATGTTTCTTTGTTGAACATGCTCTCACATTGATCAACAACATTT 1478
DB 301 CTAGAGGCCCTGATTTTCATGTTTCTTTGTTGAACATGCTCTCACATTGATCAACAACATTT 360
QY 1479 AAAGATGAAGAAAGAAAGATCAGAGAAACCTGAAATGATGATGATGTTGGAGATTAAG 1538
DB 361 AAAGATGAAGAAAGAAAGATCAGAGAAACCTGAAATGATGATGATGTTGGAGATTAAG 420
QY 1539 AAGCAGTTGTCCAACTATGAATCTCAACTTTTCAACAAATGAGGAGAAAGTAGATACAGAT 1598
DB 421 AAGCAGTTGTCCAACTATGAATCTCAACTTTTCAACAAATGAGGAGAAAGTAGATACAGAT 480
QY 1599 GATCGAACTGAAGCTATTTACGAGCAGACTCACAAGAGCCCTCCCACTTTGATTTCTCAG 1658
DB 481 GATCGAACTGAAGCTATTTACGAGCAGACTCACAAGAGCCCTCCCACTTTGATTTCTCAG 540
QY 1659 CAGCCTCAGCTCTTGAAGAGAGAGGTCATGATAGCTCATGCTCATCCACAGGAAGTC 1718
DB 541 CAGCCTCAGCTCTTGNAGAGAGAGGTCATGATAGCTCATGCTCATCCACAGGAAGTC 600
QY 1719 TACAATGAATATGTACCCAGAGGTCGAAGAATAAATGGCATTTCACATTTCCACGATACA 1778
DB 601 TACAATGAATATGTACCCAGAGGTCGAAGAATAAATGGCATTTCACATTTCCACGATACA 660
QY 1779 CTGGCCAGTCACAGCATCTCATCCACGACCATCATGATCATATTTCTCCATCAT 1838
DB 661 CTGGCCAGTCACAGCATCTCATCCACGACCATCATGATCATATTTCTCCATCAT 720
QY 1839 CACCACCAAAACCCACCATCTCACAGTCACAGCCAGCGTACTCTCGGAGAGGCTG 1898
DB 721 CACCACCAAAACCCACCATCTCACAGTCACAGCCAGCGTACTCTCGGAGAGGCTG 780
QY 1899 AAAGATGCCGGCTGCCACTTTGGCCTGGATGGTGATAATGGGTGATGGCTGCACCAAT 1958
DB 781 AAAGATGCCGGCTGCCACTTTGGCCTGGATGGTGATAATGGGTGATGGCTGCACCAAT 840
QY 1959 TTCAGCGATGGCTAGCAATTTGGTCTGCTTTTACTGAAGGCTTATCAAGTGGTTAAAGT 2018
DB 841 TTCAGCGATGGCTAGCAATTTGGTCTGCTTTTACTGAAGGCTTATCAAGTGGTTAAAGT 900
QY 2019 ACTTCGTGTGTGTCTCATGATGGCTCATGATTAAGTGACTTTTCTGCTCTTA 2078
DB 901 ACTTCGTGTGTGTCTCATGATGGCTCATGATTAAGTGACTTTTCTGCTCTTA 960
QY 2079 CTAAGGCTGGCATGACCGTTAAGCAGGCTGTCTTTTAAATGCAATGTCAGCCATGCTG 2138
DB 961 CTAAGGCTGGCATGACCGTTAAGCAGGCTGTCTTTTAAATGCAATGTCAGCCATGCTG 1020
QY 2139 GCGTATCTTGGATGGCAACAGGAATTTTCATTGCTATTTGCTGAAATGTTTCTATG 2198
DB 1021 GCGTATCTTGGATGGCAACAGGAATTTTCATTGCTATTTGCTGAAATGTTTCTATG 1080
QY 2199 TGGATATTGGCACTTACTGCTGGCTTATTCATGATGTTGCTGTTGATATGTAACCT 2258
DB 1081 TGGATATTGGCACTTACTGCTGGCTTATTCATGATGTTGCTGTTGATATGTAACCT 1140
QY 2259 GAAATGCTGCACAATGATGCTAGTCACCATGATCTACCCCTGGGGGTATTTCTTTTA 2318
DB 1141 GAAATGCTGCACAATGATGCTAGTCACCATGATCTACCCCTGGGGGTATTTCTTTTA 1200

QY 2319 CAGAACTGCTGGGATGCTTTTGGGTTTGGAAATATGTTACTTATTTCCATATTTGAACAT 2378
DB 1201 CAGAACTGCTGGGATGCTTTTGGGTTTGGAAATATGTTACTTATTTCCATATTTGAACAT 1259
QY 2379 AAAATCGTCTTCGTATAAATTTCTAGTTTAAAGTTTAAATGCTAGAGTACGTTAAAAAGT 2438
DB 1260 AAAATCGTCTTCGTATAAATTTCTAGTTTAAAGTTTAAATGCTAGAGTACGTTAAAAAGT 1318
QY 2439 TGTCTATGTTTTCAGTAGTTCATAGGAGATGAGTTTGTATGCTGTACTATGAGCGTTTA 2498
DB 1319 TGTCTATGTTTTCAGTAGTTCATAGGAGATGAGTTTGTATGCTGTACTATGAGCGTTTA 1378
QY 2499 AAGTTAGTGGGTTTTCGTGATTTTGTATGAATAATGCTGTGTACAAAAGTCAAGTTAA 2558
DB 1379 AAGTTAGTGGGTTTTCGTGATTTTGTATGAATAATGCTGTGTGTACAAAAGTCAAGTTAA 1438
QY 2559 AGGTACGTTTAAATATTTAAAGTTTATCTATCTTGGAGATAAAATCTGTATGTGCAATTTCA 2618
DB 1439 AGGTACGTTTAAATATTTAAAGTTTATCTATCTTGGAGATAAAATCTGTATGTGCAATTTCA 1498
QY 2619 CCGGTATTAACAGTTTATTAATGCTAAACAGAGATTTGGCATGACATGTTCTGTATGTTTC 2678
DB 1499 CCGGTATTAACAGTTTATTAATGCTAAACAGAGATTTGGCATGACATGTTCTGTATGTTTC 1558
QY 2679 AGGGAAAAATGCTTTAAATGCTTTTCAAGAACTAACACAGTTTATCTCTACTGGAATTT 2738
DB 1559 AGGGAAAAATGCTTTAAATGCTTTTCAAGAACTAACACAGTTTATCTCTACTGGAATTT 1618
QY 2739 TAGGCTCTGAAGAACTGCTGCTGTTTAAAGAAATAAGAAATAAGAGAAAGAAATCTGAGAA 2798
DB 1619 TAGGCTCTGAAGAACTGCTGCTGTTTAAAGAAATAAGAAATAAGAGAAAGAAATCTGAGAA 1678
QY 2799 AAGAAAGCTTATCTGAAATTTAAGCAAAATAAAGAGAAAGAAAGAAATCTGAGAA 2858
DB 1679 AAGAAAGCTTATCTGAAATTTAAGCAAAATAAAGAGAAAGAAAGAAATCTGAGAA 1738
QY 2859 TTGGGAGGCACTAGATTTCTTATAAAATCACAATAATTTGTTGTAATTTAGAGGGAGAA 2918
DB 1739 TTGGGAGGCACTAGATTTCTTATAAAATCACAATAATTTGTTGTAATTTAGAGGGAGAA 1798
QY 2919 TTTAGAAATTAAGTATAAAAGGCAAGATTTAGTATAGATACATTTCAATTAACATTTTGT 2978
DB 1799 TTTAGAAATTAAGTATAAAAGGCAAGATTTAGTATAGATACATTTCAATTAACATTTTGT 1858
QY 2979 CAGATTTATTTCCCGTAAAAAGCTAGTGAGCACTCTCATATCTAAATTTAGTGTACATTTA 3038
DB 1859 CAGATTTATTTCCCGTAAAAAGCTAGTGAGCACTCTCATATCTAAATTTAGTGTACATTTA 1918
QY 3039 ACTTTGTATAATACAGAAATCTAAATATATTTAATGAATTCAGCAATATACACTTCAC 3098
DB 1919 ACTTTGTATAATACAGAAATCTAAATATATTTAATGAATTCAGCAATATACACTTCAC 1978
QY 3099 AAGAAATTTGGAATTTCAAAATGTTGCGGGTTATATACAGATGAGTACAGTGTAG 3158
DB 1979 AAGAAATTTGGAATTTCAAAATGTTGCGGGTTATATACAGATGAGTACAGTGTAG 2038
QY 3159 TTTATGTATACCACTGCGGTTATTCGCAAGTTATATATACCAAAAGCTGTATGACTG 3218
DB 2039 TTTATGTATACCACTGCGGTTATTCGCAAGTTATATATACCAAAAGCTGTATGACTG 2098
QY 3219 GATGTTCTGGTTACCTGGTTTACAAAATTTATCAGAGTAGT----- 3258
DB 2099 GATGTTCTGGTTACCTGGTTTACAAAATTTATCAGAGTAGTAAAACTTTGATATATATGAG 2158
QY 3259 ----- 3258
DB 2159 GATATTAACACTACACTAAGTATCATTTGATTCGATTCAGAAAACTTTGATATATATGAG 2218
QY 3279 GATATTAACACTACACTAAGTATCATTTGATTCGATTCAGAAAACTTTGATATATCTCTC 3338
DB 2219 GATATTAACACTACACTAAGTATCATTTGATTCGATTCAGAAAACTTTGATATATCTCTC 2278
QY 3339 AGTGCTTCAGTGTATCATTTGTGAGCAATTTGCTTTTATATACGGTACTGTAGCCACTATA 3398

QY 313 AGCTTTTCTACCGCTATGAGAAAATAATTCTTTGTGTCAGTTGAAGGTTTCAGAAAATTAC 372
DB 121 AGCTATTCTACCGCTACGGAGAGAAATGATTTCTTGTGTTGAAGGTTTCAGAAAATTGC 180
QY 373 TTCAAAATATAGGATAGATAAGATTAAAGAAATCCATATACACCATGACCCAGCAGCATC 432
DB 181 TTCAGACATAGGATAGATAAGATTAAAGAGTCCATATACACCATGACCCAGCAGCATC 240
QY 433 ACTCAGACCCAGGACATCACTCAGACCATGACCGGTCTACTCAGACCATGAGCATCACTCAG 492
DB 241 ATGCTACCCAGGACATCACTCGGACCATGAGCATCACTCGGACCATGAGCATCACTCGG 300
QY 493 ACCACGAGCATCACTCTGACCATGATCATCACTCTCACCATAATCATGCTGCTTCGGTA 552
DB 301 ACCACGAGCATCACTCGGACCATGAGCATCACTCGGACCATGAGCATCACTCGGTA 360
QY 553 AAATAAGCGAAAGCTCTTTGCCCCAGACCATGACTCAGATAGTTCAGGTAAAGATCCCTA 612
DB 361 AAACAAATCGAAAGCGCTTTGTCCAGACCTTGACTCTGATTAATTCAGGTAAATCCCTA 420
QY 613 GAAACGCCAGGGGAAGGAGCTCACCGACCAAGACATGCCAGTGTAGAGGAATGTCA 672
DB 421 GAAGTCTGAGGAAGGATCTCGGCCAGCAGAGCAGATGAATGTTAGGAGGAACATCA 480
QY 673 AGGACAGTGTAGTGTAGTGAAGTGAACCTCAACTGTGTACAACTGTCTCTGAAGGAA 732
DB 481 AGGAGTGTCAAGCTCTAGTGAAGTGAACCTCGGCGGTATACAACTGTCTCTGAAGGAA 540
QY 733 CTCACTTTCTAGAGACAAATAGAGACTTCCAAAGACCTGGAAACTCTTCCCCAAAGATGTAA 792
DB 541 CTGCGCTTGTAGAGACAAATAGAGACTTCCAAACCTGGGAGAC---GCACAAAGATGTAA 597
QY 793 GCAGCTCCACTCCACCCAGTGTACATCAAGAGCCGGGTGAGCCGGCTG-----G 843
DB 598 AC-CTTCTACCCACCCAGCATCAGGAGAAAGCCGAGTGGCCGGCTGAGTCCGGCTAG 656
QY 844 CTGGTAGGAAACAAATGAATCTGTGAGTGAAGCCCGGAAAGGCTTTATGTATTCAGAA 903
DB 657 CTAGGAAGAAAGCAATGAGTCTGTGAGTGAAGCCAGAAAGGCTTTATGTATTCAGAA 716
QY 904 ACACAAATGAATCTCAGGAGTGTTCATATGCATCAAGCTACTGACATCTCATGCGCA 963
DB 717 ACACAAATGACAAATTCAGGAGTGTTCATATCAACCAAGCTGTGACATCCCATGCGCA 776
QY 964 TGGCATCCAGGTTCCGCTCAATGCAACAGACTTCAACTATCTCTGTCCAGCCATCATCA 1023
DB 777 TGAGCATCCAGGCTCTGTTGAATGCAACGGAATTTAACTATCTCTGCCAGCCATCATCA 836
QY 1024 ACCAAATTCATGCTAGATCTTGTCTGTGATTCATA---CAAGTGAAGAGGCTGAAATCC 1080
DB 837 ATCAAAATTCATGCTCGGGCTTCTCTGATTCATACAGCAAGTGAAGAGGCGAGAAATCC 896
QY 1081 CTCCTAAGACCTATTCAATFACAAATAGCCTGGGTTGGTGTGTTATAGCCATTTCCATCA 1140
DB 897 CTCCTAAGACCTATTCTTTACAAATAGCCTGGCTGGTGGTTCATAGCCATTTCCATCA 956
QY 1141 TCAGTTTCTGCTCTGCTGCGGGGTATCTTATAGTCCCTCTCATGAATCGGTTGTTTTCA 1200
DB 957 TCAGTTTCTGCTCTGCTGCGGGGTATCTTATAGTCCCTCTCATGAACCGGGTATTTTCA 1016
QY 1201 AATTTCTCTGAGTTTCTTGTGGCAGCTGGCGTTGGGACCTTGTAGTGGTGTATGCTTTTT 1260
DB 1017 AGTTCTGCTGAGCTTCTCTGCTGCGGGCTGGCGTGGCAACGCTGAGTGGCGATGCTCTGT 1076
QY 1261 TACACCTTCTCCACATCTCTCATGCAAGTCAACCACTAGTCAATGACCATGAAGAACAG 1320
DB 1077 TACATCTCTCCACACTCTCATGCAAGTCAATGACCATGAGTCAATGACCATGAGGACCG 1136
QY 1321 CAATGGAATGAAAGAGGACCACTTTTTCAGTCACTCTCTCTCTCAAAACATAGAAGAAA 1380
DB 1137 CGATGGAAATGAAAGAGGCCCCCTGTTTCAGCCACCTGTCGGCTCAGAAATATAGAAGAA 1196

QY 1381 GTGCTTATTTTGTATTCACAGGTGGAAGGGTCTAACAGCTCTAGGAGGCTGTATTTTCATGT 1440
DB 1197 GCTCTTATTTTGTATTCACAGGTGGAAGGGTCTACAGGCTCTAGGAGGCTTATATTTTCATGT 1256
QY 1441 TTCTTTGTTGAACATCTCTCACATTTGATCAACAATTTAAAGATAAGAAAAGAAATC 1500
DB 1257 TTCTTTGTTGAACATCTCTCACATTTGATCAACAATTTAAAGATAAGAAAAGAAATC 1316
QY 1501 AGAAGAAACCTGAAATGATGATGATGTTGAGATTTAAGAAGCAGTTTGTCCAAGTATGAAT 1560
DB 1317 AAAGAAACCTGAAATGATGAGGATGTTGAGAGCAAGAACGCTGTCCAATATCAGACT 1376
QY 1561 CTCACATTTTCAACAATGAGGAGAAAGTAGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1377 CTCACATTTTCCATGAAAGAGAGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1436
QY 1621 GAGCAGACTCAAGAGAGGCTCCCACTTTTGTATTCACAGGCTTCCAGCTTCTTGAAGAG 1680
DB 1437 GAGCAGACTCCAGAGAGGCTCCCACTTTTGTATTCACAGGCTTCCAGCTTCTTGAAGAGG 1496
QY 1681 AAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1497 AAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1556
QY 1741 GGTGCAAGAATAAATGAGGATTCACATTTCCAGAGTACACTCGGCGAGTCTAGACGATCTCA 1800
DB 1557 GCTGCAAGAATAAATGAGGATTCACATTTCCAGAGTACACTCGGCGAGTCTAGACGATCTCA 1616
QY 1801 TTCACACCATCATGACTTACCATCATATTTCCATCATCATCACCAACCAACCAACCAACCA 1860
DB 1617 TCCACACCATCATGACTTACCATCATATTTCCATCATCATCACCAACCAACCAACCAACCA 1676
QY 1861 CTCACAGTCAAGGAGGCTTACTCTCGGAGGAGCTGAAAGATGCGGCGGCTCGCCACTT 1920
DB 1677 CTCACAGGCAAGGAGGCTTACTCTCGAGAGGAGCTGAAAGAGTCCCGGCTTCCACAT 1736
QY 1921 TGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
DB 1737 TGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1796
QY 1981 GTGCTGCTTTTACTGAAAGGCTTATCAAGTGTGTTTAAAGTACTTCTGTTGCTGTGTTGCTC 2040
DB 1797 GTGCTGCTTTTCAAGGAGGCTTGTCCAGTGGCTTAAAGCCTCTGCTGCTGTGTTGCTGCTC 1856
QY 2041 ATGAGTTTCCCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
DB 1857 ATGAGTTTCCCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1916
QY 2101 AGCAGGCTGTCTTTTATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
DB 1917 AGCAGGCTGTCTTTTATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1976
QY 2161 GAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
DB 1977 GAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2036
QY 2221 GCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
DB 2037 GCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2096
QY 2281 GTGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
DB 2097 GTGATCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2156
QY 2341 GTTTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
DB 2157 GTTTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2216
QY 2401 TCTAGTTTAAAGTTTAAATGCTAGAGTGTAAAGTGTCTCATGATTTTCAGTGTAGTGTAT 2460
DB 2217 TCTAATTTAGTTTCTAAACCCCTAGAGGAGTGTAAAGAGTACTGTTGTTGTTGTTGTTGTTGTT 2276
QY 2461 AG-----GGAGTGTAGTTTGTATGCTGTACTATGACGCGTTTAAAGTTAGTGGGTTTTCT 2515

Db 2277 AGATGGTAGAAATACGTTTCATATGCTGTGATACACAGCATTTAAACTAGTGGTTTGC 2336
QY 2516 GATTTTGTATCAATATGCTCTCTGTT-----ACAAAGTCAGT 2555
Db 2337 AATTTTCTATCAAAATTTTGGCCATTTGTTATGCTTATAAGTCAGTCCATTGGGTAACA 2396
QY 2556 TAAAGGTACGTTTAAATATTAAGTTATTCATCTTGTGGAGATAAATCTGTATGTCGAAT 2615
Db 2397 TAAAGGTACGTTTAAATATTAAGTTATTCATCTTGTGGAGATAAATCTGTATGTCGAAT 2456
QY 2616 TCACCCGCTATTACCACTGTTTATATGTAACAAGAGATTTGGCATGACATGTTCTGTATGT 2675
Db 2457 CCACCACTATTACCCAGCTGT-----ATGTTTCCGCATGCCGTGCTCTCTAT 2505
QY 2676 TTCAGGGAATAATGCTTTAAATGCTTTTCAAGAACTACACATGTTATCTTATACCTGA 2735
Db 2506 TCTAGAG- AAATGCTTTAGTCTCTCTGTGTAGCTTTAGCATTTACTTCCCTGCTGTA 2564
QY 2736 TTTTAGTCTCTGAAGAACTGCTGTGTGTAGGAATAAGATGTGCATGAAGCCCTAAAT 2795
Db 2565 GTTGGGCCCT-----TGCTGCTGTGTGAGAAAGATGGSCACAAACCCAGAC 2612
QY 2796 ACCAAGAAAGCTTATACGAAATTAAGCAAAAGAAATAAGGAGAAAGAGAAATCTGA 2855
Db 2613 ACCAAGCTTACACTGAGCAT-----AGTCAGGAAATAGGAGAAAGAGAACTCGGAG 2668
QY 2856 GAATTTGGGAGCATAGATCTTTATAAAATCACAATAATTTGTTAAATAGAG-GGGA 2914
Db 2669 CATTTGAGGAGGTATATATGCTTTTGAATAATCACAAGGTGCTGTTAACTTTGAGAGGGG 2728
QY 2915 GAAATTTAGAAATTAAGTATAAAGGCGAGAATTTAGTATAGATACATTCATTAACAATTT 2974
Db 2729 GAAATTTAGATGTTTAAAGGGGAGATGTTTATGTCAGAGTACGTCGCTGGACATCTG 2788
QY 2975 TTGTGAGGATTTATTCCTGTAACCAAGCTAGTGAGCACTCTCATATACTAATAGTGACA 3034
Db 2789 TCAAAATCTTATGTTCTCTGTGAACA--AGTGAACACTCTGTATACAC---TTCATCTGCT 2843
QY 3035 TTTAACTTTGATTAATACAGAAATCTAAATATATTTAAATGAATTCAGCAATATACACTT 3094
Db 2844 TGTACATCTGTATGATAGAAATGTAATATGTT-----AATCCAAGCAATGTCACAT 2898
QY 3095 GACCAAGAAATTTGAAATTTCAAAATGTTCTGTGGGGTTATATACCAGATCAGTACAGTA 3154
Db 2899 TGGCTGAGAAATGAGGCAATTTAAATCTGTGTGG-----AGATGGGCCAAATGA 2948
QY 3155 GTAGTTTATGTATCACCC---AGACTGGGTTATTTGCCAAGTTATATATACCAAAAGCTGT 3211
Db 2949 GTAACTAGTGTGTACCAGAAAGACTTGGTTTTCACCATGCTCTCTTATGCCCTGAAGTTAT 3008
QY 3212 ATGACTGGATGTTCTGTTACCTGTTTACAAAATATACAGAGTAGTAAACTTTGATAT 3271
Db 3009 GTGA-----ACATTTCTAGTGTCTAGGCTTCCAAAGT--GGCAAGTATTAACACTTTAGCCC 3062
QY 3272 ATATGAGGATPATAAAC-----TACACTAAGTATCATTTGATTTGATTTCCAGAAAGTACT 3326
Db 3063 ATGAGAGGATGTTAAGACCTGTTTCTGAAATATTTTGACTTCTTGGAATTAAGAAAGCACT 3122
QY 3327 TGCATATCTCTAGTCTCTCAGTGCTA-----TCATTTGAGCAAAATGTC--TTTATATA 3379
Db 3123 TAGCTCTCTTCTCTCTCAGTCTGTTTAAATATGTCATTTGTGAACAAGTCTCTCTTGTATGTA 3182
QY 3380 CGGTACTGTAGGCATACTAGGCTGTCTGTGGCATCTCTCTACATGCTTTCTTTTACACA 3439
Db 3183 TGACACTGTAGATACATTTGGGCTGTCTGACACTGTCTGATGTTTCTTTTAAATA 3242
QY 3440 ATAAATTCCTTATATCAG 3457
Db 3243 ATTAAGTGTCAACCCAG 3260

RESULT 8

AX017261
LOCUS AX017261
DEFINITION Sequence 12 from Patent W09947669.
ACCESSION AX017261
VERSION AX017261.1 GI:10042179
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1597)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pillarsky,C.
TITLE Human nucleic acid sequences from tissue of breast tumors
JOURNAL Patent: WO 9947669-A 12-23-SEP-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILLARSKY CHRISTIAN (DE)
FEATURES
source Location/Qualifiers
1..1597
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 476 a 229 c 337 g 555 t
ORIGIN
Query Match 43.0%; Score 1486.8; DB 6; Length 1597;
Best local similarity 99.0%; Pred. No. 0;
Matches 1559; Conservative 0; Mismatches 7; Indels 8; Gaps 6;
QY 1895 GCTGAAGATGCCGGCTGCCACTTTGGCTGGATGGTGATGAATGGGTGATGCCCTGCA 1954
Db 19 GCTGAAGATGCCGGCTGCCACTCTGGCTGGATGGTGATGAATGGGTGATGCCCTGCA 78
QY 1955 CAATTTACAGGATGGCTAGCAATTTGGTCTCTTTTACTGAAGCTTATCAAGTGGTTT 2014
Db 79 CAATTTACAGGATGGCTAGCAATTTGGTCTCTTTTACTGAAGCTTATCAAGTGGTTT 138
QY 2015 AAGTACTTCTGTTGCTGTCTTCTGTCATGAGTTGGCTCATGAATAGTGGTCTTCTGCT 2074
Db 139 AAGTACTTCTGTTGCTGTCTTCTGTCATGAGTTGGCTCATGAATAGTGGTCTTCTGCT 198
QY 2075 TCTACTAAAGGCTGGCATGACCGTTAAGCAGGCTGTCTTTTATAATGCAATGTCAGCAT 2134
Db 199 TCTACTAAAGGCTGGCATGACCGTTAAGCAGGCTGTCTTTTATAATGCAATGTCAGCAT 258
QY 2135 GCTGGGATCTTGAATGGCAACAGGAATTTTCATTTGCTCATTTATGCTGAAAATGTTTC 2194
Db 259 GCTGGGATCTTGAATGGCAACAGGAATTTTCATTTGCTCATTTATGCTGAAAATGTTTC 318
QY 2195 TATGTGGATATTTGCATCTACTGCTGGCTTATTCATGATGTTGCTCTGGTGTGATGCT 2254
Db 319 TATGTGGATATTTGCATCTACTGCTGGCTTATTCATGATGTTGCTCTGGTGTGATGCT 378
QY 2255 ACCTGAAATGCTGCAATGATGCTAGTACCATGGATGATAGCCGCTGGGCTATTTCTT 2314
Db 379 ACCTGAAATGCTGCAATGATGCTAGTACCATGGATGATAGCCGCTGGGCTATTTCTT 438
QY 2315 TTTACAGAACTGCTGGATGCTTTTGGGTTTGGAAATATGTTACTTATTTCCATATTTGA 2374
Db 439 TTTACAGAACTGCTGGATGCTTTTGGGTTTGGAAATATGTTACTTATTTCCATATTTGA 498
QY 2375 ACATAAAATCGTGTCTGATATAAATTTCTAGTAAAGCTTTAAATGCTAGAGTACCTAAA 2434
Db 499 ACATAAAATCGTGTCTGATATAAATTTCTAGTAAAGCTTTAAATGCTAGAGTACCTAAA 558
QY 2435 AAGTTGTCATAGTTTTCAGTAGGTCATAGGGAGATGAGTTTGTATGCTGTACTATGCAAGC 2494
Db 559 AAGTTGTCATAGTTTTCAGTAGGTCATAGGGAGATGAGTTTGTATGCTGTACTATGCAAGC 618
QY 2495 TTTAAAGTTAGTGGTTTCTGTTGATTTTGTATGAAATATGCTGTCTGTCTACAAAGTCAG 2554
Db 619 TTTAAAGTTAGTGGTTTCTGTTGATTTTGTATGAAATATGCTGTCTGTCTACAAAGTCAG 678

QY	2555	TTAAAGGTACGTTTAAATATTAAAGTTATTTCTATCTTGGAGATAAAATCTGTATGCGAA	2614
Db	679	TTAAAGGTACGTTTAAATATTAAAGTTATTTCTATCTTGGAGATAAAATCTGTATGCGAA	738
QY	2615	TTCAACCGGTATTAACCAAGTATTATGTAACCAAGAGATTGGCATGACATGTTCTGTATG	2674
Db	739	TTCAACCGGTATTAACCAAGTATTATGTAACCAAGAGATTGGCATGACATGTTCTGTATG	798
QY	2675	TTTCAGGGAAATAATGCTTTTAATGCTTTTTCAGGAACATAACACAGTTATTCCTATACATGG	2734
Db	799	TTTCAGGGAAATAATGCTTTTAATGCTTTTTCAGGAACATAACACAGTTATTCCTATACATGG	858
QY	2735	ATTTTATGCTCTCAAGAACTGCTGCTGTTTAGGAATAGAAATGTCATGAAGCCTAAAA	2794
Db	859	ATTTTATGCTCTCAAGAACTGCTGCTGTTTAGGAATAGAAATGTCATGAAGCCTAAAA	918
QY	2795	TACCAAGAAAGCTTATACAGAAATTAAGCAAGAAATAAGGAGAGAAAGAGAGAAATCTG	2854
Db	919	TACCAAGAAAGCTTATACAGAAATTAAGCAAGAAATAAGGAGAGAAAGAGAGAAATCTG	978
QY	2855	AGAAATGGGAGGCATAGATTCTTATAAAATCAAAAATTTGGTGTAAATTAGAGGGGA	2914
Db	979	AGAAATGGGAGGCATAGATTCTTATAAAATCAAAAATTTGGTGTAAATTAGAGGGGA	1038
QY	2915	GAAATTTAGAATTAAGTATATAAAGCGCAGAAATAGTATAGAGTACATTATTAACATTT	2974
Db	1039	GAAATTTAGAATTAAGTATATAAAGCGCAGAAATAGTATAGAGTACATTATTAACATTT	1098
QY	2975	TTGTCAGGATTTATTCGCGTAAACAGTAGTGAG--CACTCTCATATACTAAAT--AGTG	3030
Db	1099	TTGTCAGGATTTATTCGCGTAAACAGTAGTGAG--CACTCTCATATACTAAAT--AGTG	1158
QY	3031	TACATTTAACTTTGTATATAACAGAAATCTAAATATATTTAATGAATTCAGCAATATA-	3089
Db	1159	TACATTTAACTTTGTATATAACAGAAATCTAAATATATTTAATGAATTCAGCAATATA	1218
QY	3090	CACTTTACCAGAAATTCGAAATTTCAAAATGTCGTCGGGTTATATACCAGATGAGTAC	3149
Db	1219	CACTTTACCAGAAATTCGAAATTTCAAAATGTCGTCGGGTTATATACCAGATGAGTAC	3177
QY	3150	AGTGAGTAG-TTTATGTATACACAGACTGGGTTATTCGCAAGTTATATATACCAAAAGC	3208
Db	1278	AGTGAGTAGTTTATGTATACACAGACTGGGTTATTCGCAAGTTATATATACCAAAAGC	1337
QY	3209	TGATGACTGGATGTCGTTTACCTGGTTACAAAATATACAGATAGTAAACATTTGA	3268
Db	1338	TGATGACTGGATGTCGTTTACCTGGTTACAAAATATATACAGATAGTAAACATTTGA	1397
QY	3269	TATATATCAGGATATATAAACTACATTAAGTATTCGATTCGATTCAGAAAGTACTTTT	3328
Db	1398	TATATATCAGGATATATAAACTACATTAAGTATTCGATTCGATTCAGAAAGTACTTTT	1457
QY	3329	GATATCTCTCAGTCGTCAGTGGTATCATTTGTGAGCAATGTCTTTTATATACGGTACTG	3387
Db	1458	GATATCTCTCAGTCGTCAGTGGTATCATTTGTGAGCAATGTCTTTTATATACGGTACTG	1517
QY	3388	TAGCCATCTAGGCGTCTGTGGCATCTCTAGATGTTCTTTTTTACACAATAAATTC	3447
Db	1518	TAGCCATCTAGGCGTCTGTGGCATCTCTAGATGTTCTTTTTTACACAATAAATTC	1577
QY	3448	CTTATATACGCTTG	3461
Db	1578	CTTATATACGCTTG	1591

[illegible]

ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 1310) Manning,D,Lockwood., Green,C,Douglas., Nicholson,R,Ian., Gee,J,Margaret,Wendy., and
TITLE	Methods for predicting the behaviour of breast tumours
JOURNAL	Patent: US 5693465-A 1 02-DEC-1997;
FEATURES	Location/Qualifiers 1. .1310 /organism="unknown"
BASE COUNT	354 a 271 c 287 g 398 t
ORIGIN	

Query Match					
Best Local Similarity		37.2%;	Score 1288;	DB 6;	Length 1310;
Matches 1310; Conservative		99.8%;	Pred. No. 3.5e-279;	Mismatches 0;	Indels 2; Gaps 2;
QY	1119	GGTTTTATAGCCCATTTCCCATCATCAGTTCCTTGCTCTGCTCGCTGGGGTTATCTTAGTGCCT	1178		
Db	1	GGTTTTATAGCCCATTTCCCATCATCAGTTCCTTGCTCTGCTCGCTGGGGTTATCTTAGTGCCT	60		
QY	1179	CTCATGAATCGGGTGTTTTTCAAAATTTCTCCTGAGTTTCCTTGCTGGCACTGGCGGTGGG	1238		
Db	61	CTCATGAATCGGGTGTTTTTCAAAATTTCTCCTGAGTTTCCTTGCTGGCACTGGCGGTGGG	120		
QY	1239	ACTTTGAGTGGTGATGCTTTTTACACCTTTCTCCACATTTCTCCACATTTCTGCAAGTCACCACAT	1298		
Db	121	ACTTTGAGTGGTGATGCTTTTTACACCTTTCTCCACATTTCTCCACATTTCTGCAAGTCACCACAT	180		
QY	1299	AGTCATAGCCCATGAAGAACCAGCAATGGAATGAAAGAGGAGCACACTTTTCAGTCATCTG	1358		
Db	181	AGTCATAGCCCATGAAGAACCAGCAATGGAATGAAAGAGGAGCACACTTTTCAGTCATCTG	240		
QY	1359	TCTTCTCAAACATAGAAGAAAGTGCCCTATTTTGATTCCAGTGGAAAGGGTCTTAACAGCT	1418		
Db	241	TCTTCTCAAACATAGAAGAAAGTGCCCTATTTTGATTCCAGTGGAAAGGGTCTTAACAGCT	300		
QY	1419	CTAGGAGGCCCTGATTTTCATGTTTCTTGCTGAACATGTCTTCACATTTGATCAACAATTTT	1478		
Db	301	CTAGGAGGCCCTGATTTTCATGTTTCTTGCTGAACATGTCTTCACATTTGATCAACAATTTT	360		
QY	1479	AAAGATAAAGAAGAAAAGAAATCAGAGAAACCTGAAATGATGATGTGGAGATTAAAG	1538		
Db	361	AAAGATAAAGAAGAAAAGAAATCAGAGAAACCTGAAATGATGATGTGGAGATTAAAG	420		
QY	1539	AAGCAGTTGTCCAAGTATGAATCTCAACTTTTCAACAAATGAGGAGAAAGTAGATACAGAT	1598		
Db	421	AAGCAGTTGTCCAAGTATGAATCTCAACTTTTCAACAAATGAGGAGAAAGTAGATACAGAT	480		
QY	1599	GATCCAACCTGAAGGCTATTTACGAGCAGACATCAAGAGGCCCTCCCACCTTTGATTTCTCAG	1658		
Db	481	GATCGNACTGAAGGCTATTTACGAGCAGACATCAAGAGGCCCTCCCACCTTTGATTTCTCAG	540		
QY	1659	CAGCCTGCAGTCTTTGGAGAGAAAGAGGTCTATGATAGCTCATCTCATCCACAGGAAGTCT	1718		
Db	541	CAGCCTGCAGTCTTTGGAGAGAAAGAGGTCTATGATAGCTCATCTCATCCACAGGAAGTCT	600		
QY	1719	TACAAATGAATATGTACCCAGAGGGTGCAGAATAAATGCGATTCACATTTCCACGATACA	1778		
Db	601	TACAAATGAATATGTACCCAGAGGGTGCAGAATAAATGCGATTCACATTTCCACGATACA	660		
QY	1779	CTCGGCCAGTCAGACGATCTCATTTCCACACCATCATGACTACCATCATATTTCTCCATCAT	1838		
Db	661	CTCGGCCAGTCAGACGATCTCATTTCCACACCATCATGACTACCATCATATTTCTCCATCAT	720		
QY	1839	CACCACCACCAAAACACCATCCTCACATGCACGCCAGCGCTACTCTCGGAGGAGCTG	1898		
Db	721	CACCACCACCAAAACACCATCCTCACATGCACGCCAGCGCTACTCTCGGAGGAGCTG	780		
QY	1899	AAAGATCCCGGCTCGGCCACTTTGGGCTTGSATGGGTGATTAATGGGTGATGGCCTGCACAA	1958		
Db	781	AAAGATCCCGGCTCGGCCACTTTGGGCTTGSATGGGTGATTAATGGGTGATGGCCTGCACAA	840		


```
OY 1959 TTCACGGATGGCCTAGCAATTTGGTCTGCTTTTACTGAAGGCTTATCAAGTGGTTAAAGT 2018
Db 841 TTCACGGATGGCCTAGCAATTTGGTCTGCTTTTACTGAAGGCTTATCAAGTGGTTAAAGT 900
OY 2019 ACTTCTGTTGCTGCTTCTCTCATGAGTTGCCTCATGAATAGTGACTTTGCTGTTCTTA 2078
Db 901 ACTTCTGTTGCTGCTTCTCTCATGAGTTGCCTCATGAATAGTGACTTTGCTGTTCTTA 960
OY 2079 CTAAGGCTGGCATCAGCCTTTAAGCAGGCTGCTCTTTATATGATGCTGAGCAATGCTG 2138
Db 961 CTAAGGCTGGCATCAGCCTTTAAGCAGGCTGCTCTTTATATGATGCTGAGCAATGCTG 1020
OY 2139 GCGTATCTTGAATGGCAACAGGAATTTTCATTGGTCATTATGCTGAAATGTTTCTATG 2198
Db 1021 GCGTATCTTGAATGGCAACAGGAATTTTCATTGGTCATTATGCTGAAATGTTTCTATG 1080
OY 2199 TGGATATTTGCATTTACTGCTGCTTATTCATGATGCTGCTGCTGCTGCTGCTGCTGCT 2258
Db 1081 TGGATATTTGCATTTACTGCTGCTTATTCATGATGCTGCTGCTGCTGCTGCTGCTGCT 1140
OY 2259 GAAATGCTGCACATGATGCTAGTACCATGATGATGATGATGATGATGATGATGATGAT 2318
Db 1141 GAAATGCTGCACATGATGCTAGTACCATGATGATGATGATGATGATGATGATGATGAT 1200
OY 2319 CAGATGCTGGGATGCTTTGGGCTTTTGGGCTTTTGGGCTTTTGGGCTTTTGGGCTTT 2378
Db 1201 CAGATGCTGGGATGCTTTGGGCTTTTGGGCTTTTGGGCTTTTGGGCTTTTGGGCTTT 1259
OY 2379 AAAATGCTGTTTCTGATAAATTTCTAGTTAAAGTTTAAATGCTAGAGTAGCT 2430
Db 1260 AAAATGCTG-TTCGTATAATTTCTAGTTAAAGTTTAAATGCTAGAGTAGCT 1310

RESULT 10
AC091060/c
LOCUS
DEFINITION Homo sapiens chromosome 18, clone RP11-723J4, linear PRI 28-NOV-2001
ACCESSION AC091060
VERSION AC091060.10 GI:16646826
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 221941)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Camarata, J., Campoliano, A., Chang, J., Choepel, Y., Colangelo, M.,
Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K.,
Diaz, J.S., Dodge, S., Faro, S., Ferrelra, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glnde, S., Goyette, M., Graham, L., Hulme, W.,
Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Sever, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (28-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 4, 2001 this sequence version replaced gi.16118168.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L12269
Center clone name: 723-J_4
-----
Location/Qualifiers
1. .221941
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-723J4"
/clone_lib="RPC1-11 Human Male BAC"
240. .261
/notes="<30 qual SNGL region"
322. .326
/notes="<30 qual SNGL region"
635. 670
/repeat_region
/repeat_family="(TTTA)n"
complement(674. .963)
/repeat_family="AluSg"
688. .715
/notes="<30 qual SNGL region"
951. .955
/notes="<30 qual SNGL region"
1010. .1013
/notes="<30 qual SNGL region"
1124. .1441
/repeat_region
/repeat_family="AluSx"
2833. .2926
/repeat_family="(CA)n"
complement(2960. .3257)
/repeat_family="AluJo"
complement(3291. .3371)
/repeat_family="L2"
3807. .4209
/repeat_family="L1MD1"
complement(4298. .4597)
/repeat_family="AluSx"
```

REFERENCE
AUTHORS

3 (bases 1 to 221941)

Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A.,
Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferrelra, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Glnde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Sever, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (28-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 4, 2001 this sequence version replaced gi.16118168.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L12269

Center clone name: 723-J_4

FEATURES
source

unsure
unsure
repeat_region
repeat_region
unsure
unsure
unsure
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

```
repeat_region 5434..5597
repeat_region /rpt_family="FRAM"
repeat_region complement(6415..6580)
repeat_region /rpt_family="MIR"
repeat_region complement(7324..7451)
repeat_region /rpt_family="MLT1H"
repeat_region 7505..7750
repeat_region /rpt_family="AluJb"
repeat_region complement(7768..7984)
repeat_region /rpt_family="MLT1H"
repeat_region 8022..8284
repeat_region /rpt_family="MLT1J"
repeat_region complement(8865..9153)
repeat_region /rpt_family="AluSc"
repeat_region 9584..9680
repeat_region /rpt_family="AluSg/x"
repeat_region 11358..11524
repeat_region /rpt_family="AluSx"
repeat_region 11555..11586
repeat_region /rpt_family="(TAAA)n"
repeat_region 11587..11750
repeat_region /rpt_family="L1ME2"
repeat_region 11838..12141
repeat_region /rpt_family="AluSx"
repeat_region complement(12195..12762)
repeat_region /rpt_family="MLT1E2"
repeat_region 12824..12950
repeat_region /rpt_family="AluSg/x"
repeat_region 12955..13254
repeat_region /rpt_family="AluSx"
repeat_region 13256..13920
repeat_region /rpt_family="L1MB2"
repeat_region 13921..14181
repeat_region /rpt_family="AluSx"
repeat_region 14182..14516
repeat_region /rpt_family="L1MB2"
repeat_region 14572..14886
repeat_region /rpt_family="L1ME2"
repeat_region 14905..14936
repeat_region /rpt_family="MER58B"
repeat_region 14937..15226
repeat_region /rpt_family="AluSg"
repeat_region 15227..15376
repeat_region /rpt_family="MER58B"
repeat_region 15377..16032
repeat_region /rpt_family="L1ME2"
repeat_region complement(16039..16172)
repeat_region /rpt_family="L1MCA"
repeat_region complement(16197..17024)
repeat_region /rpt_family="L1MCA"
repeat_region complement(17025..17305)
repeat_region /rpt_family="AluSx"
repeat_region complement(17306..17710)
repeat_region /rpt_family="L1MCA"
repeat_region 17710..18046
repeat_region /rpt_family="L1ME2"
repeat_region 18047..18330
repeat_region /rpt_family="GA-rich"
repeat_region 18336..19013
repeat_region /rpt_family="L1ME2"
repeat_region 19014..19286
repeat_region /rpt_family="AluYa5"
repeat_region 19287..19895
repeat_region /rpt_family="L1ME2"
repeat_region 19896..20066
repeat_region /rpt_family="AluJb"
repeat_region 20070..20562
repeat_region /rpt_family="L1ME2"
repeat_region 20680..20722
repeat_region /rpt_family="L1ME2"
repeat_region 20817..20961
repeat_region /rpt_family="L1MC5"
repeat_region complement(20990..21142)
```

```
repeat_region /rpt_family="L1MC5"
repeat_region 21143..21430
repeat_region complement(21431..21443)
repeat_region /rpt_family="L1MC5"
repeat_region 21444..21466
repeat_region /rpt_family="(TAAA)n"
repeat_region complement(21467..21637)
repeat_region /rpt_family="L1MC5"
repeat_region complement(21654..21943)
repeat_region /rpt_family="AluJb"
repeat_region complement(22021..22069)
repeat_region /rpt_family="L1MA8"
repeat_region 22070..22092
repeat_region /rpt_family="(TA)n"
repeat_region complement(22093..22227)
```

Query Match 32.9% Score 1137.8; DB 9; Length 221941;
Best Local Similarity 98.9%; Pred. No. 3.5e-245;
Matches 1209; Conservative 0; Mismatches 7; Indels 7; Gaps 6;

```
QY 2245 TTGATATGGTACCTGAATGCTGCACAAATGATGCTAGTGACCATGATGATGACCGCTGGG 2304
Db 170112 TTTTAAGGTACTGAAATGCTGCACAAATGATGCTAGTGACCATGATGATGACCGCTGGG 170053
QY 2305 GGTATTCTTTTACAGAAATGCTGGGATGCTTTTGGGTTTGGGAATTATGTTACTATT 2364
Db 170052 GGTATTCTTTTACAGAAATGCTGGGATGCTTTTGGGTTTGGGAATTATGTTACTATT 169993
QY 2365 CCATATTTGAACATAAATCGTGTTCGTATAAATTTCTAGTTAAGGTTTAAATGCTAGA 2424
Db 169992 CCATATTTGAACATAAATCGTGTTCGTATAAATTTCTAGTTAAGGTTTAAATGCTAGA 169933
QY 2425 GTAGCTTAAAAAGTTGTCATAGTTTTCAGTAGTATAGGAGATGAGTGTGATGCTGTA 2484
Db 169932 GTAGCTTAAAAAGTTGTCATAGTTTTCAGTAGTATAGGAGATGAGTGTGATGCTGTA 169873
QY 2485 CTATGCAGGTTTAAAGTTAGTGGGTTTGTGATTTTGTATTTGTAATTTCTGCTGCTT 2544
Db 169872 CTATGCAGGTTTAAAGTTAGTGGGTTTGTGATTTTGTATTTGTAATTTCTGCTGCTT 169813
QY 2545 ACAAAGTCAGTTAAAGGTACGTTTAAATTTAAAGTTATTTCTATCTTGGAGATAAATCT 2604
Db 169812 ACAAAGTCAGTTAAAGGTACGTTTAAATTTAAAGTTATTTCTATCTTGGAGATAAATCT 169753
QY 2605 GTATGTGCAATTCACCGGTATTACCGTTTATTATGTAACAAGAGATTTGGCATGACAT 2664
Db 169752 GTATGTGCAATTCACCGGTATTACCGTTTATTATGTAACAAGAGATTTGGCATGACAT 169693
QY 2665 GTTCTGTATGTTTCAGGGAAAAATGCTTTAATGCTTTTCAAGAACTAACACAGTTATT 2724
Db 169692 GTTCTGTATGTTTCAGGGAAAAATGCTTTAATGCTTTTCAAGAACTAACACAGTTATT 169633
QY 2725 CCTATCTGGATTTTAGTCTCTGAAGAATGCTGGTGTGTTTAGGAATAAGAAATGTCATG 2784
Db 169632 CCTATCTGGATTTTAGTCTCTGAAGAATGCTGGTGTGTTTAGGAATAAGAAATGTCATG 169573
QY 2785 AAGCCTTAAATACCAACAAGCTTTACTACTGAATTTAAGCAAGAAATTAAGGAGAAAAA 2844
Db 169572 AAGCCTTAAATACCAACAAGCTTTACTACTGAATTTAAGCAAGAAATTAAGGAGAAAAA 169513
QY 2845 GAAGAATCTCAGAAATTTGGGGAGGCATAGATTCTTATAAATAACCAAAATTTGTTGTAAA 2904
Db 169512 GAAGAATCTCAGAAATTTGGGGAGGCATAGATTCTTATAAATAACCAAAATTTGTTGTAAA 169453
QY 2905 TTAGAGGGGAGAAATTTAGAAATTAAGTATAAAGGCAGAAATTTAGATAGTACATTCA 2964
Db 169452 TTAGAGGGGAGAAATTTAGAAATTAAGTATAAAGGCAGAAATTTAGATAGTACATTCA 169393
QY 2965 TTAACATTTTTCAGGATTATTTCCCGTAAAAACGTTAGTGAGCAC-TCTCATATACTA 3023
Db 169392 TTAACATTTTTCAGGATTATTTCCCGTAAAAACGTTAGTGAGCACTTTTCATATACTA 169333
```

QY 3024 ATT--AGTGACATTAACTTTGTATAATACAGAAATCTAAATATATATTTAAATGAATCAA 3081
 Db 169332 ATTTAGTTGTACATTTAACTTTGTATAATACAGAAATCTAAATATATATTTAAATGAATCAA 169273
 QY 3082 GCAATATA-CACTTGACCAAGAAATGGAAATTCAGGATGTCGCGGGTATATATACCA 3140
 Db 169272 GCATATATACACTTGACCAAGAAATGGAAATTCAGGATGTCGCGGGTATATATACCA 169214
 QY 3141 GATGACTACAGTAGTAG-TTTATGTATACACAGACTGGGTTATTTGCAAGTTATATATAT 3199
 Db 169213 GATGACTACAGTAGTAGTTTATGTATACACAGACTGGGTTATTTGCAAGTTATATATAT 169154
 QY 3200 ACCAAAGCTGTATGACAGTGGATGTCGTTACCTGCTGTTTACAAAATTTATCAGAGTAGTA 3259
 Db 169153 ACCAAAGCTGTATGACAGTGGATGTCGTTACCTGCTGTTTACAAAATTTATCAGAGTAGTA 169094
 QY 3260 AAATTTGATATATATGAGGATATATAAACTACACTAAGTATCATTTGATTCGATTCAGA 3319
 Db 169093 AAATTTGATATATATGAGGATATATAAACTACACTAAGTATCATTTGATTCGATTCAGA 169034
 QY 3320 AAGTACTTTGATATCTCTCAGTCTCAGTCTCAGTCTATCATTTGAGCAATGTC-TTTATAT 3378
 Db 169033 AAGTACTTTGATATCTCTCAGTCTCAGTCTCAGTCTATCATTTGAGCAATGTC-TTTATAT 168974
 QY 3379 ACGGTACTGAGGCATCTAGGCTCTGCTGTCGATTCCTAGATGTTCTTTTATATAT 3438
 Db 168973 ACGGTACTGAGGCATCTAGGCTCTGCTGTCGATTCCTAGATGTTCTTTTATATAT 168914
 QY 3439 AATAAAATTCCTTATATACAGCTTG 3461
 Db 168913 AATAAAATTCCTTATATACAGCTTG 168891

RESULT 11

AP001905

LOCUS

DEFINITION

AP001905 Homo sapiens chromosome 18 clone RP11-723J4 map 18q12, WORKING

DRAFT SEQUENCE, 23 unordered pieces.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

* NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 42444: contig of 42444 bp in length
 * 42445 42544: gap of 100 bp
 * 42545 64659: contig of 22115 bp in length
 * 64660 64759: gap of 100 bp
 * 64760 82876: contig of 18117 bp in length
 * 82877 82976: gap of 100 bp
 * 82977 100906: contig of 17930 bp in length
 * 100907 101006: gap of 100 bp
 * 101007 117951: contig of 16945 bp in length
 * 117952 118051: gap of 100 bp
 * 118052 131503: contig of 13452 bp in length
 * 131504 131603: gap of 100 bp
 * 131604 141537: contig of 9934 bp in length
 * 141538 152889: contig of 11252 bp in length
 * 152890 152989: gap of 100 bp
 * 152990 161818: contig of 8829 bp in length
 * 161819 161918: gap of 100 bp
 * 161919 170983: contig of 9065 bp in length
 * 170984 171083: gap of 100 bp
 * 171084 178239: contig of 7156 bp in length
 * 178240 178339: gap of 100 bp
 * 178340 187572: contig of 9233 bp in length
 * 187573 187672: gap of 100 bp
 * 187673 193032: contig of 5360 bp in length
 * 193033 193132: gap of 100 bp
 * 193133 198842: contig of 5710 bp in length
 * 198843 198942: gap of 100 bp
 * 198943 203032: contig of 4090 bp in length
 * 203033 203132: gap of 100 bp

Consensus quality: 219708 bases at least Q30
 Consensus quality: 221297 bases at least Q20
 Insert size: 222588; sum-of-contigs
 Quality coverage: 9.33x in Q20 bases; sum-of-contigs

 NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 42444 contig of 42444 bp in length
 42545 64659 contig of 22115 bp in length
 82876 82976 contig of 18117 bp in length
 100906 101007 contig of 17930 bp in length
 117951 118051 contig of 16945 bp in length
 131503 131604 contig of 13452 bp in length
 141537 141638 contig of 9934 bp in length
 152889 152990 contig of 11252 bp in length
 161818 161919 contig of 8829 bp in length
 170983 171084 contig of 9065 bp in length
 178239 178340 contig of 7156 bp in length
 187572 187673 contig of 9233 bp in length
 193032 193133 contig of 5360 bp in length
 198842 198943 contig of 5710 bp in length
 203032 203133 contig of 4090 bp in length
 207783 207884 contig of 4651 bp in length
 211428 211529 contig of 3545 bp in length
 215591 215692 contig of 4063 bp in length
 217832 217933 contig of 2141 bp in length
 220510 220611 contig of 2578 bp in length
 222035 222136 contig of 1425 bp in length
 223655 223756 contig of 1520 bp in length
 224788 contig of 1033 bp in length

* NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 42444: contig of 42444 bp in length
 * 42445 42544: gap of 100 bp
 * 42545 64659: contig of 22115 bp in length
 * 64660 64759: gap of 100 bp
 * 64760 82876: contig of 18117 bp in length
 * 82877 82976: gap of 100 bp
 * 82977 100906: contig of 17930 bp in length
 * 100907 101006: gap of 100 bp
 * 101007 117951: contig of 16945 bp in length
 * 117952 118051: gap of 100 bp
 * 118052 131503: contig of 13452 bp in length
 * 131504 131603: gap of 100 bp
 * 131604 141537: contig of 9934 bp in length
 * 141538 152889: contig of 11252 bp in length
 * 152890 152989: gap of 100 bp
 * 152990 161818: contig of 8829 bp in length
 * 161819 161918: gap of 100 bp
 * 161919 170983: contig of 9065 bp in length
 * 170984 171083: gap of 100 bp
 * 171084 178239: contig of 7156 bp in length
 * 178240 178339: gap of 100 bp
 * 178340 187572: contig of 9233 bp in length
 * 187573 187672: gap of 100 bp
 * 187673 193032: contig of 5360 bp in length
 * 193033 193132: gap of 100 bp
 * 193133 198842: contig of 5710 bp in length
 * 198843 198942: gap of 100 bp
 * 198943 203032: contig of 4090 bp in length
 * 203033 203132: gap of 100 bp

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: <http://hgp.gsc.riken.go.jp/>

Contact: hattori@hgp.gsc.riken.go.jp

----- Project Information

Center project name: HumDrafit18

Center clone name: RP11-723J4

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 216741 bases at least Q40


```

OY 3320 AGTACTTTGATATCTCTCAGTGGCTTCACTGCTATCATGTGAGCAATGTC-TTTATAT 3378
|||||
Db 20682 AACTACTTTGATATCTCTCAGTGGCTTCACTGCTATCATGTGAGCAATGTC-TTTATAT 20741
|||||
OY 3379 ACGGTACTCTAGCCATACCTAGCCCTGCTGTGGCATTCTCTAGATGTTCTTTTACAC 3438
|||||
Db 20742 ACGGTACTCTAGCCATACCTAGCCCTGCTGTGGCATTCTCTAGATGTTCTTTTACAC 20801
|||||
OY 3439 AATAAATCTTATATCAGCTTGG 3461
|||||
Db 20802 AATAAATCTTATATCAGCTTGG 20824
|||||

RESULT 12
AP001158
LOCUS
DEFINITION Homo sapiens chromosome 18 clone RP11-701C9 map 18q12, WORKING
DRAFT SEQUENCE, 15 unordered pieces.
ACCESSION AP001158
VERSION AP001158.3 GI:9188478
KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-701C9.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 160170)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 160,170 genomic DNA of 18q12
Published Only in Database (2000)
2 (bases 1 to 160170)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (10-FEB-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8535, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
On Jul 14, 2000 this sequence version replaced gi:8118726.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft18
Center clone name: RP11-701C9
----- Summary Statistics
Sequencing Vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 155820 bases at least Q40
Consensus quality: 157459 bases at least Q30
Consensus quality: 158091 bases at least Q20
Insert size: 158770; sum-of-contigs
Quality coverage: 10.41x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
15 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 29590 contig of 29590 bp in length
29691 49906 contig of 20216 bp in length
50007 69120 contig of 19114 bp in length
69221 84192 contig of 14972 bp in length
84293 99890 contig of 15598 bp in length
99991 110467 contig of 10477 bp in length
110568 118491 contig of 7924 bp in length

NOTE: This is a 'working draft' sequence. It currently
consists of 15 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 29590: contig of 29590 bp in length
* 29591 29690: gap of 100 bp
* 29691 49906: contig of 20216 bp in length
* 49907 50006: gap of 100 bp
* 50007 69120: contig of 19114 bp in length
* 69121 69220: gap of 100 bp
* 69221 84192: contig of 14972 bp in length
* 84193 84292: gap of 100 bp
* 84293 99890: contig of 15598 bp in length
* 99891 99990: gap of 100 bp
* 99991 110467: contig of 10477 bp in length
* 110468 110567: gap of 100 bp
* 110568 118491: contig of 7924 bp in length
* 118492 118591: gap of 100 bp
* 118592 126449: contig of 7858 bp in length
* 126450 126549: gap of 100 bp
* 126550 134958: contig of 8409 bp in length
* 134959 135058: gap of 100 bp
* 135059 141053: contig of 5995 bp in length
* 141054 141153: gap of 100 bp
* 141154 145920: contig of 4767 bp in length
* 145921 146020: gap of 100 bp
* 146021 149439: contig of 3419 bp in length
* 149440 149539: gap of 100 bp
* 149540 154173: contig of 4634 bp in length
* 154174 154273: gap of 100 bp
* 154274 157616: contig of 3343 bp in length
* 157617 157716: gap of 100 bp
* 157717 160170: contig of 2454 bp in length.

FEATURES
Source
1..160170
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18q12"
/clone="RP11-701C9"

misc_feature
1..29590
/note="assembly_fragment"
29691..49906
/note="assembly_fragment clone_end:T7 vector_side:right"
50007..69120
/note="assembly_fragment"
69221..84192
/note="assembly_fragment"
84293..99890
/note="assembly_fragment"
99991..110467
/note="assembly_fragment"
110568..118491
/note="assembly_fragment clone_end:sp6 vector_side:right"
118592..126449
/note="assembly_fragment"
126550..134958
/note="assembly_fragment"
135059..141053
/note="assembly_fragment"
141154..145920
/note="assembly_fragment"

```

BASE COUNT	46219 a	33232 c	33436 g	45883 t	1400 others
ORIGIN	/note= assembly_fragment"				

Matches 836; Conservative 0; Mismatches 115; Indels 3; Gaps 2;

QY . 3082 GCAATATACACTTGACCAAGAAATTGGAATTCAAATGTTCTGTCGGGTTATATACCAG 3141

[illegible]

[illegible]

Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shooshcarl, N., Sisson, I.,
 Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 193437)
 Worley, K.C.
 Direct Submission
 Submitted (08-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 193437)
 Worley, K.C.
 Direct Submission
 Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 10, 2002 this sequence version replaced gi:21703549.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GBAI
 Center clone name: CH230-5P17
 ----- Summary Statistics
 Sequencing vector: Plasmid:
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 143517 bases at least Q40
 Consensus quality: 150307 bases at least Q30
 Consensus quality: 154257 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 69 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 1172: contig of 1172 bp in length
 * 1173 1272: gap of unknown length
 * 1273 2487: contig of 1215 bp in length
 * 2488 2587: gap of unknown length
 * 2588 4394: contig of 1807 bp in length
 * 4395 4494: gap of unknown length
 * 4495 5815: contig of 1321 bp in length
 * 5816 5915: gap of unknown length
 * 5916 7626: contig of 1711 bp in length
 * 7627 9207: gap of unknown length
 * 9207 9306: contig of 1480 bp in length
 * 9307 10449: contig of 1143 bp in length
 * 10450 10549: gap of unknown length
 * 10550 11609: contig of 1060 bp in length
 * 11610 11709: gap of unknown length
 * 11710 12957: contig of 1248 bp in length
 * 12958 13057: gap of unknown length
 * 13058 14501: contig of 1444 bp in length
 * 14502 14601: gap of unknown length
 *
 * 14602 15762: contig of 1161 bp in length
 * 15763 15862: gap of unknown length
 * 15863 16877: contig of 1015 bp in length
 * 16878 16977: gap of unknown length
 * 16978 18761: contig of 1784 bp in length
 * 18762 18861: gap of unknown length
 * 18862 20537: contig of 1676 bp in length
 * 20538 22730: contig of 2093 bp in length
 * 22731 22830: gap of unknown length
 * 22831 25262: contig of 2432 bp in length
 * 25263 27262: gap of unknown length
 * 27263 27362: contig of 1900 bp in length
 * 27363 28524: gap of unknown length
 * 28525 28624: gap of unknown length
 * 28625 30233: contig of 1609 bp in length
 * 30234 30334: gap of unknown length
 * 30335 31883: contig of 1550 bp in length
 * 31884 31984: gap of unknown length
 * 31985 33437: contig of 1453 bp in length
 * 33438 34740: gap of unknown length
 * 34741 34839: gap of unknown length
 * 34840 36917: contig of 2078 bp in length
 * 36918 37017: gap of unknown length
 * 37018 38173: contig of 1156 bp in length
 * 38174 39552: contig of 1679 bp in length
 * 39553 40053: gap of unknown length
 * 40054 42001: contig of 1948 bp in length
 * 42002 43702: contig of 1602 bp in length
 * 43703 43802: gap of unknown length
 * 43803 46149: contig of 2347 bp in length
 * 46150 46249: gap of unknown length
 * 46250 47598: contig of 1349 bp in length
 * 47599 47698: gap of unknown length
 * 47699 49211: contig of 1512 bp in length
 * 49212 49310: gap of unknown length
 * 49311 50780: contig of 1470 bp in length
 * 50781 50880: gap of unknown length
 * 50881 52850: contig of 1970 bp in length
 * 52851 52950: gap of unknown length
 * 52951 53981: contig of 1031 bp in length
 * 53982 54081: gap of unknown length
 * 54082 55821: contig of 1740 bp in length
 * 55822 55921: gap of unknown length
 * 55922 57909: contig of 1988 bp in length
 * 57910 58009: gap of unknown length
 * 58010 60388: contig of 2379 bp in length
 * 60389 60488: gap of unknown length
 * 60489 62460: contig of 1972 bp in length
 * 62461 62560: gap of unknown length
 * 62561 64500: contig of 1940 bp in length
 * 64501 64600: gap of unknown length
 * 64601 66883: contig of 2289 bp in length
 * 66884 69589: gap of unknown length
 * 69590 69689: gap of unknown length
 * 69690 71653: contig of 1964 bp in length
 * 71654 71753: gap of unknown length
 * 71754 75273: contig of 3520 bp in length
 * 75274 75374: gap of unknown length
 * 75375 77772: contig of 2399 bp in length
 * 77773 77872: gap of unknown length
 * 77873 81413: contig of 3541 bp in length
 * 81414 81513: gap of unknown length
 * 81514 85249: contig of 3736 bp in length
 * 85250 85349: gap of unknown length
 * 85350 87366: contig of 2017 bp in length
 * 87367 87466: gap of unknown length
 * 87467 90125: contig of 2659 bp in length

TITLE
JOURNAL

REFERENCE
AUTHORS

TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Db 108242 ATTCCAGAACCTCAAATGACAAATATCCAGGAG 108273

Search completed: November 8, 2002, 19:49:16
Job time : 7252.03 secs

Query Match		12.4%;	Score 429.2;	DB 2;	Length 193437;
Best Local Similarity		75.7%;	Pred. No. 1.1e-85;		
Matches 615;		Conservative 0;	Mismatches 128;	Indels 69;	Gaps 4;
QY 127	ACGAAGCGCAATGGCGAGGAAGTATCTGTAACTTTGATCTCTGACCTTTGGCCCTCTCTG 186				
Db 107519	AGGAAGGTGCAATGGCCACAAATTTATCTGTATCATGATCTTGACCTTTGCCCTTTGGG 107578				
QY 187	TCACAAATCCCTTCATGAACATAA---AGCAGCTGCTTTCCCCAGACCACTGAGAAAA 243				
Db 107579	TTACAAATCCCTTCATGAACATAAATCAACAGCTGCTTTCTCTCAGACCACTGAGAAAA 107638				
QY 244	TTAGTCGGAATGGGAATCTGGCATTAAATCTTGACTTGGCAATTTCCACACGSCAATATC 303				
Db 107639	TTAATTCAAATTCGGGAATCTGGGATTAATGTTGACCTGGCAGTTACCATGCGAGCGACATC 107698				
QY 304	ATCTACACAGCTTTTCTACCGCTATGGAGAAATAATTTTGTGTCAGTTGAAGGGTTCA 363				
Db 107699	ATCTGCAACAGCTTTTCTACCGCTATGGAGAAATGACTCTCTGTGTCAGTTGAAGGGTTCA 107758				
QY 364	GAATTTACTTTCAAAATATAGGCATAGATTAAGATTAAAGAAATCCATATACACCATGACC 423				
Db 107759	GAATTTACTTTCAAAATATAGGCATAGATTAAGATTAAAGAAATCCATATACACCATGACC 107818				
QY 424	ACGACCATCACTCAGACCCAGGAGCATCACTCAGACCATGAGGTCACCTCAGACCATGAGC 483				
Db 107819	AC-----GAGC 107824				
QY 484	ATCACTCAGACCAAGGAGCATCACTCTGACCATGATCATCACTCTCACCATAATCATGCTG 543				
Db 107825	GCCACTCTGACCAAGGAGGCGCACTCTGACCAAGGAGGCGCCACTCCCAAGGAGGTCATGCTG 107884				
QY 544	CTTCTGGTAAAAATAAGCGAAAGCTCTTTGCCAGACCACTGACTCAGATAGTTCAAGGTA 603				
Db 107885	CCGCTGGGAAAAACAGTCGGAAAGCCTTTTGCCAGACCTTGACTCTGATAATTCAGGTA 107944				
QY 604	AAGATCTAGAACACAGCGAGGAAAGGAGCTCACCAGCAGCAACATGCCAGTGGTAGAA 663				
Db 107945	AAATCGGAACACACCGCAAGGTAAAGGATCTCGCCACAGAGCATGTGAATGGCAGGA 108004				
QY 664	GGAATGTCAAGGACAGTGTAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 723				
Db 108005	GGAATGGAAGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 108064				
QY 724	CTGAAGGAACCTCACTTTCTAGACACAATAGAGACTCCAGACCTGGAAGAACTCTTCCCA 783				
Db 108065	CTGAAGGAACCTCACTTTCTAGACACAATAGAGACTCCAGACCTGGAAGAACTCTTCCCA 108121				
QY 784	AAGATGTAGCAGCTCCACTCCACCCAGTGTACATCAAGAGCGGGGTGAGCGGCTG- 842				
Db 108122	AAGATATAAACCTTCTACCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 108181				
QY 843	-----GCTGTAGGAAACAAATGAATCTGTGTAGTGAAGGAGGAGGAGGAGGAGGAGGAG 894				
Db 108182	GCGGACTAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 108241				
QY 895	ATTCCAGAACACAAATGAAATCTTCAGGAG 926				

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2002, 17:37:02 ; Search time 3433.14 Seconds
(without alignments)
16326.890 Million cell updates/sec

Title: US-09-642-034-1
Perfect score: 3461
Sequence: 1 ctctgfcgaattcgacg.....aaattccttatcagcttg 3461

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: gb_gss:*
 - 18: em_gss_hum:*
 - 19: em_gss_inv:*
 - 20: em_gss_pln:*
 - 21: em_gss_vrt:*
 - 22: em_gss_fun:*
 - 23: em_gss_nam:*
 - 24: em_gss_mus:*
 - 25: em_gss_other:*
 - 26: em_gss_pro:*
 - 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	877.6	25.4	1074	13	BM480018
2	800.6	23.1	915	14	BQ716456
3	799.6	23.1	872	9	AU120027
4	738.8	21.3	1059	13	BM557996
5	724	20.9	810	13	BM456317
6	711	20.5	712	14	BM721841

7	694	20.1	782	12	BG571701
8	692	20.0	701	14	BQ027619
9	676.8	19.6	728	12	BG403446
10	661.6	19.1	830	13	B1222452
11	652.6	18.9	697	9	AL039253
12	652	18.8	1135	13	BM560789
13	649.8	18.8	674	14	BM978669
14	633.2	18.3	673	9	A1744245
15	627.6	18.1	957	12	BG168169
16	625.2	18.1	720	14	BQ013982
17	617.2	17.8	927	14	BQ926045
18	613.4	17.7	791	12	BG547794
19	605.4	17.5	746	12	BG530601
20	597.6	17.3	604	14	BQ581865
21	591.8	17.1	857	12	BF032013
22	591.4	17.1	743	9	A1907176
23	586.8	17.0	766	14	BM948026
24	576	16.6	925	12	BF983458
25	571.4	16.5	581	9	A1635070
26	567.4	16.4	829	9	AU079709
27	566.8	16.4	751	14	BM963737
28	566.2	16.4	887	13	B1150324
29	564	16.3	781	14	BM949686
30	561.4	16.2	684	14	BQ009520
31	560	16.2	670	14	BM999466
32	559	16.2	559	10	AW771378
33	555.2	16.0	708	12	BF978788
34	551.8	15.9	765	12	BG168944
35	538.6	15.6	948	14	BQ955836
36	538.4	15.6	733	10	BE391889
37	526.6	15.2	580	12	BF915495
38	523.6	15.1	836	12	BF790668
39	517.8	15.0	849	12	BF680209
40	510.4	14.7	797	12	BF679305
41	505.2	14.6	682	14	BM949679
42	505.2	14.6	834	12	BF669257
43	503.4	14.5	950	14	BQ433766
44	502.2	14.5	518	10	AW178411
45	501.8	14.5	513	10	AW178409

ALIGNMENTS

RESULT 1
BM480018
LOCUS
DEFINITION AGENCOURT_6424145 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5491572
5', mRNA sequence.
ACCESSION BM480018
VERSION BM480018.1 GI:18529060
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1074)
AUTHORS NTH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM12112 row: e column: 13
High quality sequence stop: 652.
Location/Qualifiers
1. .1074

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5491572"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
326 a 256 C 220 g 271 t 1 others

```

Query Match 25.4%; Score 877.6; DB 13; Length 1074;
Best Local Similarity 94.4%; Pred. No. 1.4e-182;
Matches 954; Conservative 0; Mismatches 50; Indels 7; Gaps 4;

[illegible]

QY	1393	ATTCCAGTGGGAAGGGTCTTAACAGCTCTAGGAGCCGTGTATTTTCATGTTTCTTGTGTAAC	1452
Db	843	ATTTCGCGTGGAA-GGTCTAAGAGCTCTAGAAGCCGCGCATTTTCTGTTCCTGTTCAAC	901
QY	1453	ATCTCCTCAGATGATCAAAACAATTTAAAGATAAGAGAAGAAAAGAAATCAGAGAAACCTG	1512
Db	902	ATCCCCCCCCATTGGTC-AACCCCTTTTGGTAGGAAGAAAAGAAATCCTCAAAAACCTG	960
QY	1513	AAA-ATCATGATCATGTGGAGATTAAGAGCAGCTGTGTCCAAGTATGAA	1559
Db	961	AAAATCGATGATGATGTGGAAATTTAAAAAACCGTTTCCCCCAATGAA	1011
RESULT 2			
LOCUS	B0716456		
DEFINITION		915 bp mRNA linear EST 16-JUL-2002	
ACCESSION	AGENCOURT_8291338	Lupski_sympathetic.trunk	Homo sapiens
VERSION	IMAGE:6194809.5	mRNA sequence.	
KEYWORDS	B0716456.1	GI:21855353	
SOURCE	EST.	human.	
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 915)		
JOURNAL	NH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. James R. Lupski CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM13600 row: k column: 02 High quality sequence stop: 590. Location/Qualifiers 1 source		
FEATURES			

FEATURES
source

```

0.7213
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6194809"
/clone_lib="Lupski_sympathetic_trunk"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DRI08"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCCAGCCAGCGCCG-3' and
5'-GACTAGTTCTTAGATCGCAGCGCCGCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."
BASE COUNT      274 a 118 c 204 g 317 t 2 others
ORIGIN

Query Match      23.1%; Score 800.6; DB 14; Length 915;
Best Local Similarity 96.8%; Pred. No. 1.3e-165;
Matches 880; Conservative 0; Mismatches 21; Indels 8; Gaps 6;

QY 2040 CATGAGTTGGCCTCATGAATTAGGTGACTTTGCTGTCTACTAAAGCGTGGCATGACCGTT 2099
|||||
Db 1 CATGAGTTGCCTCATGAATTAGGTGACTTTGCTGTCTACTAAAGCGTGGCATGACCGTT 60
|||||

QY 2100 AAGCAGCGTGTCTTTATAATGCATGTTCAGCCATGCTGGCGTATCTTGGTAATGCCAACA 2159
|||||

```


JOURNAL COMMENT	Unpublished (1999)
Contact: Robert Strausberg, Ph.D.	
Email: cgapbs-re@mail.nih.gov	
Tissue procurement: Lou Staudt	
CDNA Library preparation: Life Technologies, Inc.	
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
DNA Sequencing by: Agencourt Bioscience Corporation	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
Plate: LLAM12125	row: j column: 07
High quality sequence stop: 640.	
Location/Qualifiers	
source	1..810
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:5496678"
	/clone_lib="NIH_MGC_85"
	/tissue_type="lymphoma, cell line"
	/lab_host="DH10B (phage-resistant)"
	/notes="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI; cDNA Library prepared unidirectionally; oligo-dT primed. Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies Note: this is a NIH_MGC Library."
BASE COUNT	246 a 172 c 165 g 226 t
ORIGIN	1 others
Query Match	20.9%; Score 724; DB 13; Length 810;
Best Local Similarity	96.2%; Pred. No. 9.2e-149;
Matches 764; Conservative	0; Mismatches 25; Indels 5; Gaps
QY 1003	ATCTCTGTCACGCCATCATCAACCAAAATTGATGCTAGATCTTTGTCTGATTCATCAAGTG 1062
Db	
1	ATCTCTGTCACGCCATCATCAACCAAAATTGATGCTAGATCTTTGTCTGATTCATCAAGTG 60
QY 1063	AAAGAAGGCTGAATCCCTCAAGAGCCTATTTCATTACAATAGCCCTGGGTGGTGTT 1122
Db	
61	AAAGAAGGCTGAATCCCTCAAGAGCCTATTTCATTACAATAGCCCTGGGTGGTGTT 120
QY 1123	TTATAGCCATTTCCATCATCACTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 1182
Db	
121	TTATAGCCATTTCCATCATCACTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 180
QY 1183	TGAATCGGTGTTTTTCAAAATTTCTCTGAGTTTCTTGTGTCACCTGCGCGTTGGGACTT 1242
Db	
181	TGAATCGGTGTTTTTCAAAATTTCTCTGAGTTTCTTGTGTCACCTGCGCGTTGGGACTT 240
QY 1243	TGAGTGGTGATGCTTTTTTTACACCTTTCTTCCACATCTCATGCAAGTCAACCACTAGTC 1302
Db	
241	TGAGTGGTGATGCTTTTTTTACACCTTTCTTCCACATCTCATGCAAGTCAACCACTAGTC 300
QY 1303	ATAGCCATGAAGAACCGAGCAATGGAATGAAAGAGGACCACTTTTCAGTCATCTGTCTT 1362
Db	
301	ATAGCCATGAAGAACCGAGCAATGGAATGAAAGAGGACCACTTTTCAGTCATCTGTCTT 360
QY 1363	CTCAAAACATAGAAAGAGTGCCTATTTTATTCCACGTGGAGGCTCTTAACAGCTCTAG 1422
Db	
361	CTCAAAACATAGAAAGAGTGCCTATTTTATTCCACGTGGAGGCTCTTAACAGCTCTAG 420
QY 1423	GAGGCTGTATTTTCATGTTCTTGTGAACATGTCTCTACATTTGATGATCAACAATTTAAG 1482
Db	
421	GAGGCTGTATTTTCATGTTCTTGTGAACATGTCTCTACATTTGATGATCAACAATTTAAG 480
QY 1483	ATAAGAAGAAAAGCAATCAGAAGAAACCTCAAAATGATGATGTGGAGATTAAAGAGC 1542
Db	
481	ATAAGAAGAAAAGCAATCAGAAGAAACCTCAAAATGATGATGTGGAGATTAAAGAGC 540
QY 1543	AGTTGTCCAAGTATGAATCTCAACTTTCAACAAATGAGGAGAAAGTAGATACAGATGATC 1602
Db	
541	AGTTGTCCAAGTATGAATCTCAACTTTCAACAAATGAGGAGAAAGTAGATACAGATGATC 600
QY 1603	GAACCTGAAGGCTATTTAGAGCAGACTCAAGAGCCCTCCCACTTTGATTTCTCAGCAGC 1662

Tissue Procurement: CLONTECH Laboratories, Inc. cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLC1575 row: k column: 05 High quality sequence stop: 710.									
FEATURES									
source									
1..782									
/organism="Homo sapiens"									
/db_xref="taxon:9606"									
/clone="IMAGE:4720636"									
/clone_lib="NIH_MGC_79"									
/lab_host="DH10B (T1 phage-resistant)"									
/note="Organ: placenta; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgctcgcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGGCGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."									
BASE COUNT 218 a 103 c 176 g 285 t									
ORIGIN									
Query Match 20.1%; Score 694; DB 12; Length 782; Best Local Similarity 96.9%; Pred. No. 3.7e-142; Matches 762; Conservative 0; Mismatches 15; Indels 9; Gaps 5;									
QY	2055	GAATTAGTGACTTGGCTTCTTACTAAAGCTGGCATGACCGTTAAGCAGGCTGTCCTT	2114						
Db	1	GAATTAGTGACTTGGCTTCTTACTAAAGCTGGCATGACCGTTAAGCAGGCTGTCCTT	60						
QY	2115	TATAATGATTTGCAGCCATGCTGGCGTATCTTGGAAATGGCAACAGGAATTTTCATGGT	2174						
Db	61	TATAATGATTTGCAGCCATGCTGGCGTATCTTGGAAATGGCAACAGGAATTTTCATGGT	120						
QY	2175	CATTATGCTGAAGATGTTTCTATGTTGGATATTTGCACATTTACTGCTGGCTTATTCATGTAT	2234						
Db	121	CATTATGCTGAAGATGTTTCTATGTTGGATATTTGCACATTTACTGCTGGCTTATTCATGTAT	180						
QY	2235	GTGCTCTGGTTGATATGGTACCTGAAATGCTGCACAAATGATGCTAGTGACCATGGATGT	2294						
Db	181	GTGCTCTGGTTGATATGGTACCTGAAATGCTGCACAAATGATGCTAGTGACCATGGATGT	240						
QY	2295	AGCCGCTGGGGTATTTCTTTTACAGAAATGCTGGGATGCTTTGGGTTTGGAAATATG	2354						
Db	241	AGCCGCTGGGGTATTTCTTTTACAGAAATGCTGGGATGCTTTGGGTTTGGAAATATG	298						
QY	2355	TTTACTTATTTCCATATTTTGAACATAAAATCGTGTTCGTATATAATTTCTAGTTAAGGTTT	2414						
Db	299	TTTACTTATTTCCATATTTTGAACATAAAATCGTGTTCGTATATAATTTCTAGTTAAGGTTT	358						
QY	2415	AAATGCTAGAGTAGCTTAAAAAGTTGTCATAGTTTTCAGTAGCTCATAGGAGATGAGCTTT	2474						
Db	359	AAATGCTAGAGTAGCTTAAAAAGTTGTCATAGTTTTCAGTAGCTCATAGGAGATGAGCTTT	418						
QY	2475	GTATGCTGACTATGCAGCGTTTAAAGTAGTGGGTTTGTGATTTTGTATTGAATATT	2534						
Db	419	GTATGCTGACTATGCAGCGTTTAAAGTAGTGGGTTTGTGATTTTGTATTGAATATT	476						
QY	2535	GCTGCTCTGTACAAAGTCAGTTTAAAGGTACGTTTAAATATTTAAGTTATTCATCTTGA	2594						
Db	477	GCTGCTCTGTACAAAGTCAGTTTAAAGGTACGTTTAAATATTTAAGTTATTCATCTTGA	536						
QY	2595	GATAAAATCTGATGTGCAATTCACCGGTATTACAG-TTTATATGTAAACAAGAGATT	2653						
Db	537	GATAAAATCTGATGTGCAATTCACCGGTATTACAG-TTTATATGTAAACAAGAGATT	596						

```
QY 2654 TGGCATGACATGCTCTGATGTTTCAGGAAAAATGCTTTAATGC-TTTTTCAGAACT 2712
Db 597 TGGCATGACATGCTCTGATGTTTCAGGAAAAATGCTTTAATGC-TTTTTCAGAACT 656
QY 2713 AACACAGTATTCTCTATCTACTGATTTTAGTCTCTGAGAACTGCTGGTGTTTAGGAATA 2772
Db 657 AACACAGTATTCTCTATCTACTGATTTTAGTCTCTGAGAACTGCTGGTGTTTAGGAATA 716
QY 2773 AGAATGTCATGAAGCCCTAAAAA---TACCAGAAAGCTTATCTATGAAATTTAAGCAAGAA 2829
Db 717 AGAATGTCGTTGAGGCTTAAATTTACCAGAAAGGCTTATCTATGAAATTTAAGCAAGAA 776
QY 2830 ATAAAG 2835
Db 777 ATAAAG 782

RESULT 8
LOCUS BQ027619/c
DEFINITION UI-H-CO0-aqz-a-11-0-UI.s1 NCI_CGAP_Sub9 Homo sapiens cDNA clone
IMAGE:3105668 3', mRNA sequence.
ACCESSION BQ027619
VERSION BQ027619.1 GI:19762898
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 701)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES             Location/Qualifiers
     source            1..701
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:3105668"
                        /clone_lib="NCI_CGAP_Sub9"
                        /tissue_type="mixed"
                        /dev_stage="mixed"
                        /lab_host="DH10B (Life Technologies)"
                        /note="Vector: pT73-Pac (Pharmacia) with a modified
                        polylinker; Site_1: EcoR I; Site_2: Not I; tissues:
                        Cholonic mucosa with Crohns disease, Cholonic mucosa with
                        ulcerative colitis, Fetal thymus, Cervix, Cervical
                        adenosquamous carcinoma, Ligament cells, Prostate
                        carcinoma, Bladder carcinoma, Brain oligodendrocyte;
                        NCI_CGAP_Sub9 is a subtracted cDNA library constructed
                        according to Bonaldo, Lennon and Soares, Genome Research,
                        6:791-806, 1996. First strand cDNA synthesis was primed
                        with an oligo-dT primer containing a Not I site. Double
                        stranded cDNA was ligated to an EcoR I adaptor, digested
                        with Not I, and cloned directionally into pT73-Pac
                        vector. The oligonucleotide used to prime the synthesis of
                        first-strand cDNA contains a library tag sequence that is
                        located between the Not I site and the (dT)18 tail. The
                        sequence tags for this library are CGTC, AACG, GGCGC,
                        GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional
                        information, contact: Bento Soares, Bento-soares@uiowa.edu
                        TAG_LIB=UI-H-CO0
```

```
TAG_TISSUE=Prostate Carcinoma
TAG_SEQ=ATGG*
BASE COUNT 198 a 143 c 152 g 208 t
ORIGIN
Query Match 20.0%; Score 692; DB 14; Length 701;
Best Local Similarity 99.3%; Pred. No. 1e-141;
Matches 695; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 805 CACCCAGTGTACATCAAAAGAGCCGGGTGAGCCGGTGGCTGGTAGGAAAAAATGAAT 864
Db 701 CACCCAGTGTACATCAAAAGAGCCGGGTGAGCCGGTGGCTGGTAGGAAAAAATGAAT 642
QY 865 CTGTGAGTGAGCCCGGAAAAAGGCTTTATGTATTCCAGAAACACAAAATGAAATCCCTCAGG 924
Db 641 CTGTGAGTGAGCCCGGAAAAAGGCTTTATGTATTCCAGAAACACAAAATGAAATCCCTCAGG 582
QY 925 AGTGTTCATATGATCAAAAGCTACTGACATCTCATGGCATGGGCATCCAGGTTCCGCTGA 984
Db 581 AGTGTTCATATGATCAAAAGCTACTGACATCTCATGGCATGGGCATCCAGGTTCCGCTGA 522
QY 985 ATGCAACAGAGTTCAACTATCTCTGCCAGCCATCATCAACCAAAATGTAGCTAGATCTT 1044
Db 521 ATGCAACAGAGTTCAACTATCTCTGCCAGCCATCATCAACCAAAATGTAGCTAGATCTT 462
QY 1045 GTCTGATTTCATACAAAGTGAAGAAGGCTGAAATCCCTCCAAAGACCTATTTCATTACAAA 1104
Db 461 GTCTGATTTCATACAAAGTGAAGAAGGCTGAAATCCCTCCAAAGACCTATTTCATTACAAA 402
QY 1105 TAGCCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1164
Db 401 TAGCCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 342
QY 1165 TTATCTTAGTGCCTCTCATCAATCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1224
Db 341 TTATCTTAGTGCCTCTCATGAATCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 282
QY 1225 CACTGGCCGTTGGGACCTTTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1284
Db 281 CACTGGCCGTTGGGACCTTTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 222
QY 1285 CAAGTCACACCATAGTCATAGCCATGAAGAACACAGCAATGGAATGAAAGAGGACCAC 1344
Db 221 CAAGTCACACCATAGTCATAGCCATGAAGAACACAGCAATGGAATGAAAGAGGACCAC 162
QY 1345 TTTTCAGTCATCTGCTCTCTCAAAACATAGAAGAAAGTGCCTATTTTGTATCCACGTGGA 1404
Db 161 TTTTGTAGTCATCTGCTCTCTCAAAACATAGAAGAAAGTGCCTATTTTGTATCCACGTGGA 102
QY 1405 AGGCTCTAACAGCTCTAGGAGGCGCTGTATTTCATGTTTCTTTGTGAAACATGTCTCACAT 1464
Db 101 AGGCTCTAACAGCTCTAGGAGGCGCTGTATTTCATGTTTCTTTGTGAAACATGTCTCACAT 42
QY 1465 TGATCAACAAATTTAAAGATAAGAAAGAAAGAAATCAGAA 1504
Db 41 TGATCAACAAATTTAAAGATAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2
RESULT 9
LOCUS BG403446
DEFINITION 602419134F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4526145 5',
mRNA sequence.
ACCESSION BG403446
VERSION BG403446.1 GI:13296894
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 728)
AUTHORS NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```


JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10433 row: c column: 10
High quality sequence stop: 723.
Location/Qualifiers
1. 728

FEATURES
source

19.6%; Score 676.8; DB 12; Length 728;
Best Local Similarity 98.9%; Pred. No. 2.3e-138;
Matches 724; Conservative 0; Mismatches 2; Indels 6; Gaps 4;

QY 2003 ATCAAGTGGTTAAAGTACTTCTGTTGCTGCTGCTCATGAGTGGCTCATGATTAAGG 2062
|||||
Db 1 ATCAAGTGGTTAAAGTACTTCTGTTGCTGCTGCTCATGAGTGGCTCATGATTAAGG 60
|||||

QY 2063 TGACTTTGCTGTTCTACTAAAGCGTGGCATGACCGTTAAGCAGCGTGTCCTTTAATAGC 2122
|||||

Db 61 TGACTTTGCTGTTCTACTAAAGCGTGGCATGACCGTTAAGCAGCGTGTCCTTTAATAGC 120
|||||

QY 2123 ATTGTGACGATGCTGGCGTATCTTGGAAATGCAACAGAAATTTTCATGTGCTCATTAAGC 2182
|||||

Db 121 ATTGTGACGATGCTGGCGTATCTTGGAAATGCAACAGAAATTTTCATGTGCTCATTAAGC 180
|||||

QY 2183 TGAATAATGTTCTATGTGGATATTTGCACTTACTGCTGGCTTATTCATGTAATGTGCTCT 2242
|||||

Db 181 TGAATAATGTTCTATGTGGATATTTGCACTTACTGCTGGCTTATTCATGTAATGTGCTCT 240
|||||

QY 2243 GGTTCATATGGTACCTGAAATGCTGCACAATGATGCTAGTGACCATGATGATGACCGCTG 2302
|||||

Db 241 GGTTCATATGGTACCTGAAATGCTGCACAATGATGCTAGTGACCATGATGATGACCGCTG 300
|||||

QY 2303 GGGGTATTTCTTTTACAGAATGCTGGGATGCTTTTGGGTTTGGAAATATGTTACTTAT 2362
|||||

Db 301 GGGGTATTTCTTTTACAGAATGCTGGGATGCTTTTGGGTTTGGAAATATGTTACTTAT 358
|||||

QY 2363 TTCCATATTTGAACATAAAATCGTGTTCGTAATAAATTTCTAGTTAAGGTTTAAATGCTA 2422
|||||

Db 359 TTCCATATTTGAACATAAAATCGTGTTCGTAATAAATTTCTAGTTAAGGTTTAAATGCTA 418
|||||

QY 2423 GAGTAGCTTTAAAGCTTGTCTAGTTTTCAGTAGGTCATAGGAGATGATGTTGTATGCTG 2482
|||||

Db 419 GAGTAGCTTTAAAGCTTGTCTAGTTTTCAGTAGGTCATAGGAGATGATGTTGTATGCTG 477
|||||

QY 2483 TACTATGACCGCTTTAAAGTTAGTGGGTTTGTGATTTTGTATTTGAATATGTTGCTGCTG 2542
|||||

Db 478 TACTATGACCGCTTTAAAGTTAGTGGGTTTGTGATTTTGTATTTGAATATGTTGCTGCTG 537
|||||

QY 2543 TTACAAGTCAGTTAAAGTACGTTTTTAATATTTAAGTTATTTCTATCTTGGAGATAAAAT 2602
|||||

Db 538 TTACAAGTCAGTTAAAGTACGTTTTTAATATTTAAGTTATTTCTATCTTGGAGATAAAAT 597
|||||

QY 2603 CTGTATGTCGAATTCACCGGTATTACCAGTTTATTATGTAAACAAGAGATTTGGCATGAC 2662
|||||

Db 598 CTGTATGTCGAATTCACCGGTATTACCAGTTTATTATGTAAACAAGAGATT--GGATGAC 655
|||||

QY 2663 ATGTTCTGTATGTTTCAGGG-AAAAATGCTCTTAATGCTTTTTCAGAGAACTAACACAGTT 2721
|||||

Db 656 ATGTTCTGTATGTTTCAGGGAAAAAATGCTTTAATGCTTTTTCAGAGAACTAACACAGTT 715
|||||

QY 2722 ATTCCTATACCTG 2733
|||||

Db 716 ATTCCTATACCTG 727
|||||

RESULT 10
BI222452 830 bp mRNA linear EST 11-JUL-2001
LOCUS 602940041F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5102890 5',
DEFINITION mRNA sequence.
ACCESSION BI222452
VERSION BI222452.1 GI:4675896
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 830)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11248 row: b column: 11
High quality sequence stop: 765.
Location/Qualifiers
1. 830
source /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5102890"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT 256 a 100 c 183 g 291 t

ORIGIN

Query Match 19.1%; Score 661.6; DB 13; Length 830;
Best Local Similarity 93.6%; Pred. No. 5.1e-135;
Matches 769; Conservative 0; Mismatches 39; Indels 14; Gaps 7;

QY 2186 AAATGTTTCTATGTGGATATTTGCACCTTACTGCTGGCTTATTCATGTTGCTCTGG- 2244
|||||

Db 2 AAATGTTTCTATGTGGATATTTGCACCTTACTGCTGGCTTATTCATGTTGCTCTGGC 61
|||||

QY 2245 TTGATATGGTACCTGAAATGCTGCACAATGATGCTAGTGACCATGGATGAGCCGCTGGG 2304
|||||

Db 62 TTGATATGGTACCTGAAATGCTGCACAATGATGCTAGTGACCATGGATGAGCCGCTGGG 121
|||||

QY 2305 GGTATTTCTTTTACACAATGCTGGGATGCTTTTGGGTTTGGAAATATGTTACTTATTT 2364
|||||

Db 122 GGTATTTCTTTTACACAATGCTGGGATGCTTTTGGGTTTGGAAATATGTTACTTATTT 179
|||||

QY 2365 CCATATTTGAACATAAAATCGTGTTCGTAATAAATTTCTAGTTAAGGTTTAAATGCTAGA 2424
|||||

Db 180 CCATATTTGAACATAAAATCGTGTTCGTAATAAATTTCTAGTTAAGGTTTAAATGCTAGA 239
|||||

QY 2425 GTAGCTTAAAGTTGTCATAGTTTCAGTAGGTTCATAGGAGATGAGTTGTATGCTGTA 2484
|||||
Db 240 GTAGCTTAAAGTTGTCATAGTTTCAGTAGGTTCATAGGAGATGAGTTGTATGCTGTA 299
QY 2485 CTATGACGCGTTTAAAGTTAGTGGTTTTGTGATTTTTTGTATTGAATATTGCTGCTGTT 2544
|||||
Db 300 CTATGACGCGTTTAAAGTTAGTGGTTTT -GTGATTTTGTATTGAATATTGCTGCTGTT 357
QY 2545 ACAAGTACAGTTAAAGTACGTTTAAATATTAAAGTTATTTCTCTCTGGAGATAAAATCT 2604
|||||
Db 358 ACAAGTACAGTTAAAGTACGTTTAAATATTAAAGTTATTTCTCTCTGGAGATAAAATCT 417
QY 2605 GTATGTCAAATTCACCGGTATTACAGTTTATTATGTAACAAGAGATTTTGGCATGACAT 2664
Db 418 GTATGTCAAATTCACCGGTATTACAGTTTATTATGTAACAAGAGATTTTGGCATGACAT 477
QY 2665 GTTCTGTATGTTTACAGGAAATATGCTTTAATGCTTTTTCAGAACTAACACAGTTATT 2724
|||||
Db 478 GTTCTGTATGTTTACAGGAAATATGCTTTAATGCTTTTTCAGAACTAACACAGTTATT 537
QY 2725 CCTATACGTGATTTTACGTCCTGTAAGAACTGCTGGTGTAAAGTAAGAAATGTCATG 2784
Db 538 CCTATACGTGATTTTACGTCCTGTAAGAACTGCTGGTGTAAAGTAAGAAATGTCATG 597
QY 2785 AAGCCTAAATACCAAGAAAGCTTATACCTGAATTTAAAGAAAGAAATTAAGGAGAAAGA 2844
Db 598 AAGCCTAAATACCAAGAAAGCTTATACCTGAATTTAAAGCAGAAATTAAGCAGAAAGA 657
QY 2845 GAAGAACTGAGAAATTTGGGAGGCATAGATTCCTTAT -AAAATACAAAATTTTGTGTA 2903
|||||
Db 658 GAAGAACTGAGAA -TGGGAGGCATAGATTCCTTATTAANAATCACAAAATTTGTTGTAC 716
QY 2904 ATTACAGGAGCAAAATTT -AGAAATTAAGTATAAAGAGGCAAGATTAAGTATAGAGTACA 2960
|||||
Db 717 ATTACAGGAGCAAAATTTACGAATTAAGTATTAAGAGGCGGAATTAAGTATAGATTACG 776
QY 2961 TTCATTAAACATTTTGTGAGGATTTATCCCGTAAAGACGT 3002
|||||
Db 777 TTCATTAA -TTTGCAGGATTTATCCCGTAAAGTACGT 814

RESULT 11
AL039253
LOCUS
DEFINITION DKF2p727K071_r1 727 (synonym: hmcfl) Homo sapiens cDNA clone
ACCESSION AL039253
VERSION AL039253.1 GI:5408327
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 697)
AUTHORS Poustka A., Klein M., Mewes H.W., Gassenhuber J. and Wiemann S.
TITLE EST (Poustka, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Poustka A.J.
Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innistrasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mpg-berlin-dahlem.mpg.de
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKF2p727K071) is available at the RZPD in Berlin.

please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. .697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2p727K071"
/clone_lib="727 (synonym: hmcfl)"
/tissue_type="breast cancer"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 247 a 76 c 136 g 238 t
ORIGIN
Query Match 18.9%; Score 652.6; DB 9; Length 697;
Best Local Similarity 98.9%; Pred. No. 4.8e-133;
Matches 689; Conservative 0; Mismatches 4; Indels 4; Gaps 3;
QY 2411 GTTTAAATGCTAGAGTAGCTTAAAAAGTTGTCTATAGTTTCAGTAGTCATAGGGAGATGA 2470
|||||
Db 1 GTTTAAATGCTAGAGTAGCTTAAAAAGTTGTCTATAGTTTCAGTAGTCATAGGGAGATGA 60
QY 2471 GTTTGTATGCTACTATGCGAGGTTTAAAGTTAGTGGGTTTGTGATTTTGTATTGAA 2530
|||||
Db 61 GTTTGTATGCTACTATGCGAGGTTTAAAGTTAGTGGGTTTGTGATTTTGTATTGAA 120
QY 2531 TATTGCTGCTGTACAAAGTCAGTTAAAGTCAGTTTAAATTTAAAGTTATTCCTATCT 2590
|||||
Db 121 TATTGCTGCTGTACAAAGTCAGTTAAAGTCAGTTTAAATTTAAAGTTATTCCTATCT 180
QY 2591 TGGAGATAAAATCTGTATGTGCAATTCACCGGTATTACCGAGTTTATATGTAACAAGAG 2650
Db 181 TGGAGATAAAATCTGTATGTGCAATTCACCGGTATTACCGAGTTTATATGTAACAAGAG 240
QY 2651 ATTTGGCATGACATGTTCTGTATGTTTCAGGGAATAATGCTTTTAAAGTCTTTTCAAGAA 2710
Db 241 ATTTGGCATGACATGTTCTGTATGTTTCAGGGAATAATGCTTTTAAAGTCTTTTCAAGAA 300
QY 2711 CTAAACAGTTATTCCTATAGTCTCTGAAAGAACTGCTGCTGTTTAGGAA 2770
Db 301 CTAAACAGTTATTCCTATAGTCTCTGAAAGAACTGCTGCTGTTTAGGAA 360
QY 2771 TAAGATGTGATGAAGCCCTAAATACCAAGAACTTATGCTTAAAGCAAGAA 2830
Db 361 TAAGATGTGATGAAGCCCTAAATACCAAGAACTTATGCTTAAAGCAAGAA 420
QY 2831 TAAAGGAGAAAGAGAAATCTGAGAAATGGGAGGCGATAGTTCTTATAAAATCACA 2890
Db 421 TAAAGGAGAAAGAGAAATCTGAGAAATGGGAGGCGATAGTTCTTATAAAATCACA 480
QY 2891 AAATTTGTTGTAATTTAGAGGGAGAAATTTAGAAATTTAAAGTATATAAGGAGCAAGATTAGT 2950
Db 481 AAATTTGTTGTAATTTAGAGGGAGAAATTTAGAAATTTAAAGGAGCAAGATTAGT 540
QY 2951 ATAGAGTACATTCAATTAACAACTTTTGTGAGGATTTATCCCGTAAAGAGCTAGTAGCA 3010
Db 541 ATAGAGTACATTCAATTAACAACTTTTGTGAGGATTTATCCCGTAAAGAGCTAGTAGCA 600
QY 3011 C-TCCTATATACTAATTT -AGTGTACATTTTAACTTTGTTATATAACAGAAATCTAAATATA 3067
Db 601 CTTTTCATATACTAATTTAGTTGTGACATTTAACTTTGTATATAACAGAAATCTAAATATA 660
QY 3068 TTTAATGAATTTCAAGCAATATA-CACCTTGACCAAGAA 3103
Db 661 TTTAATGAATTTCAAGCAATATACTACITGACCAAGAA 697
RESULT 12
BM560789
LOCUS
DEFINITION AGENCOURT_65660.03 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5550711
5', mRNA sequence.

Tue Nov 26 12:43:35 2002

ACCESSION	BM560789	1	GI:18805500
VERSION	EST.		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1135)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ ;		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabps@remail.nih.gov Tissue Procurement: ATCC cDNA Library preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation DNA Distribution by: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LHAM12263 row: e column: 16 High quality sequence stop: 423.		
FEATURES	Location/Qualifiers		
source	1..1135		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:550711"		
	/clone_lib="NIH_MGC_67"		
	/tissue_type="retinoblastoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: Noti; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Library constructed by Life Technologies."		
BASE COUNT	292 a 255 c 266 g 322 t		
ORIGIN			
Query Match	18.8%	Score 652;	DB 13; Length 1135;
Best Local Similarity	98.3%	Pred. No. 6.7e-133;	
Matches 680;	Conservative 0;	Mismatches 10;	Indels 2; Gaps 2;
QY 1601	TCGAACTGAAGGCTATTTACGAGCAGACTCACAGAGCGCTCCCACTTTGATTCTTCACGA 1660		
DB 2	TCTCAATGAAGGCTATTTACGAGCAGACTCACAGAGCGCTCCCACTTTGATTCTTCACGA 61		
QY 1661	GCTGCAGCTTTGGAAGAAGAAGAGGCTCATGATAGCTCATGCTCATCCACGAGGAAGTCTA 1720		
DB 62	GCTGCAGCTTTGGAAGAAGAAGAGGCTCATGATAGCTCATGCTCATCCACGAGGAAGTCTA 121		
QY 1721	CAATGAATATGTACCCAGAGGGTGCAAGATAAATGCCATTACATTTCCACGATACACT 1780		
DB 122	CAATGAATATGTACCCAGAGGGTGCAAGATAAATGCCATTACATTTCCACGATACACT 181		
QY 1781	CGCCAGCTCAGAGCTCATTTCAACCACCATCATGACTACCATCATATTTCTCCATCATCA 1840		
DB 182	CGCCAGCTCAGAGCTCATTTCAACCACCATCATGACTACCATCATATTTCTCCATCATCA 241		
QY 1841	CCACCACCAAAACCCATCCTCACAGTCACAGCCAGCGTACTCTCGGAGGAGCTGAA 1900		
DB 242	CCACCACCAAAACCCATCCTCACAGTCACAGCCAGCGTACTCTCGGAGGAGCTGAA 301		
QY 1901	AGATGCCGGGGTCCGCACCTTTGGCCCTGGATGGTAAATGGGTGATGGGCTGCACAAATTT 1960		
DB 302	AGATGCCGGGGTCCGCACCTTTGGCCCTGGATGGTAAATGGGTGATGGGCTGCACAAATTT 361		
QY 1961	CAGCGATGGCCTAGCAATTTGGTGTGCTGCTTTTACTGAAGGCTTATCAAGTGGTTAAGTAC 2020		
DB 362	CAGCGATGGCCTAGCAATTTGGTGTGCTGCTTTTACTGAAGGCTTATCAAGTGGTTAAGTAC 421		
QY 2021	TTCTGTTGCTGTGTTCTGTCTATGAGTTGCCTCATGAATTAAGTGACTTTTCTGTCTACT 2080		
DB 422	TTCTGTTGCTGTGTTCTGTCTATGAGTTGCCTCATGAATTAAGTGACTTTTCTGTCTACT 481		

Db	191	TATGTAACAAGAGATTGGCATGACATGTTCTGTATGTTTCACGGGAAAAAATGCTTTAA	132
Qy	2697	TCGTTTTCAAGAACTAACACAGTATTTCCTATACTCGATTTTGGATTTTGGTCTCTGGAAGAACTG	2756
Db	131	TGCTTTTCAAGAACTAACACAGTATTTCCTATACTCGATTTTGGTCTCTGGAAGAACTG	72
Qy	2757	CTGCTGTTTAGGATAGATGTCGATGAGCCCTAAATACCAAGAAAGCTTTTACTGAA	2816
Db	71	CTGCTGTTTAGGATAGAAATGTCATGAAGCCCTAAATACCAAGAAAGCTTTTACTGAA	12
Qy	2817	TTTAAGCAAAAG	2827
Db	11	TTTAAGCAAAAG	1
RESULT 15			
LOCUS	BG168169	957 bp	linear EST 06-FEB-2001
DEFINITION	602341563F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:449591 5', mRNA sequence.		
ACCESSION	BG168169		
VERSION	BG168169.1	GI:12674872	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 957)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10233 row: m column: 16 High quality sequence start: 14 High quality sequence stop: 659. Location/Qualifiers 1. 957 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:449591" /clone_lib="NIH_MGC_89" /tissue_type="hypernephroma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."		
BASE COUNT	321 a	229 c	204 g
ORIGIN			
Query Match	18.1%	Score 627.6;	DB 12: Length 957;
Best Local Similarity	89.6%;	Pred. No. 1.6e-127;	
Matches 731; Conservative	0; Mismatches 79; Indels 6; Gaps 5;		
Qy	425	CGACCATCTCTGACCAACAGGATCAC-TCAGACCATGAGC-GTCACCTAGACCATGAG	482
Db	1	CGACCATCTCTGACCAACAGGATCACGTCAGACCATGAGCGGTCTACTCAGAAACATGAG	60
Qy	483	C-ATCATCTGACCAACAGGATCACCTCTGACCATGATCATCATCTCTCACCATAATCATGC	541
Db	61	CAATCATCTGACCAACAGGATCATCTCTGACCATGATCATCATCTCTCACCATAATCATGC	120
Qy	542	TGC-TTCTGTTAAAAATAAGCGAAAGCTTTTGCCAGACCATGACTCATGATGATCTCAG	600

Tue Nov 26 12:43:33 2002

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2002, 17:38:12 ; Search time 96.4981 Seconds
(without alignments)
10999.262 Million cell updates/sec

Title: US-09-642-034-1
Perfect score: 3461
Sequence: 1 ctctatgcgaattcggcagc.....aaattccttatatcagcttg 3461

Scoring table: IDENTITY_NUC
 Canon 10.0 : Gapext 1.0

Searched:	441362 seqs, 153338381 residues
Total number of hits satisfying chosen parameters:	882724

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : Issued_patents_NA.*
1: /cgn2_6/ptodata/1/1na/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/1na/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/1na/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/1na/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/1na/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/1na/backfiles1.seq.*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			ID	Description
	Score	Match	Length		
1	2249	65.0	2404	1	US-08-311-023-3
2	1298	37.2	1310	1	Sequence 1, Appli
3	360	10.4	385	4	Sequence 182, Appl
c 4	165	4.8	177	4	Sequence 162, App
c 5	165	4.8	177	4	Sequence 162, App
c 6	165	4.8	177	4	Sequence 162, App
c 7	165	4.8	177	4	Sequence 162, App
c 8	165	4.8	177	4	Sequence 162, App
c 9	165	4.8	177	4	Sequence 162, App
c 10	59.8	1.7	285	4	Sequence 64, Appl
c 11	59.8	1.7	285	4	Sequence 64, Appl
c 12	59.8	1.7	342	4	Sequence 65, Appl
c 13	59.8	1.7	342	4	Sequence 65, Appl
c 14	59.8	1.7	351	4	Sequence 60, Appl
c 15	59.8	1.7	351	4	Sequence 60, Appl
c 16	59.8	1.7	356	4	Sequence 60, Appl
c 17	59.8	1.7	356	4	Sequence 60, Appl
c 18	59.8	1.7	358	4	Sequence 60, Appl
c 19	59.8	1.7	358	4	Sequence 60, Appl
c 20	59.8	1.7	363	4	Sequence 63, Appl
c 21	59.8	1.7	363	4	Sequence 63, Appl
c 22	59.8	1.7	363	4	Sequence 63, Appl
c 23	59.8	1.7	363	4	Sequence 63, Appl
c 24	59.8	1.7	409	4	Sequence 57, Appl
c 25	59.8	1.7	409	4	Sequence 57, Appl
c 26	59.8	1.7	410	4	Sequence 59, Appl
c 27	59.8	1.7	410	4	Sequence 61, Appl

us-09-642-034-1.rni

QY	2259	GAATGCTGCACAATGATGCTAGTGACCATGGATGTAGCCGCTGGGGGTATTTCTTTTTTA	2318
DB	1141	GAATGCTGCACAAATGATGCTAGTGACCATGGATGGATGGGGGTATTTCTTTTTTA	1200
QY	2319	CAGAATGCTGGGATGCTTTTGGGGTTTTCGAAATATGTTTACTTATTTCCATATTTGAACAT	2378
DB	1201	CAGAATGCTGGGATGCTTTTGGGGTTTTCGAAATATGTTTACTTATTTCCATATTTGAACAT	1259
QY	2379	AAATGCTGTTTCGTATAAATTTCTAGTTAAGGTTTAAATGCTTAGTAGTACGCTTAAAGT	2438
DB	1260	AAATGCTG-TTCGTATAAATTTCTAGTTAAGGTTTAAATGCTTAGTAGTACGCTTAAAGT	1318
QY	2439	TGTCATAGTTTCAGTAGGTCATAGGCAGATGAGTTTGTATGCTGTACTATACACGCTTTA	2498
DB	1319	TGTCATAGTTTCAGTAGGTCATAGGCAGATGAGTTTGTATGCTGTACTATACACGCTTTA	1378
QY	2499	AAGTTAGTGGGTTTGTGATTTTGTGATTTGCAATATTTGCTGCTGTACAAAGTCAGTTAA	2558
DB	1379	AAGTTAGTGGGTTTGTGATTTTGTGATTTTGTGATTTTGTGCTGCTGTACAAAGTCAGTTAA	1438
QY	2559	AGGTACGTTTAAATATTAAGTTATCTATCTTTGGAGATAAAATCTGTATGTCGAATTTCA	2618
DB	1439	AGGTACGTTTAAATATTAAGTTATCTATCTTTGGAGATAAAATCTGTATGTCGAATTTCA	1498
QY	2619	CCGGTATTACCAAGTTATTATGTAACAACAGAGATTTGGCATGACATGTTCTGTATGTTTC	2678
DB	1499	CCGGTATTACCAAGTTATTATGTAACAACAGAGATTTGGCATGACATGTTCTGTATGTTTC	1558
QY	2679	AGGGAAAAATGCTTTTAATGCTTTTCAAGAACTTAACACAGTTTATTCCTATACTCGATTT	2738
DB	1559	AGGGAAAAATGCTTTTAATGCTTTTCAAGAACTTAACACAGTTTATTCCTATACTCGATTT	1618
QY	2739	TAGGTCCTGAAGAACTGCTGGTGTTAGGAATAGAAATGTCATGAAGCCTTAAATATCC	2798
DB	1619	TAGGTCCTGAAGAACTGCTGGTGTTAGGAATAGAAATGTCATGAAGCCTTAAATATCC	1678
QY	2799	AAGAAAGCTTATACTGAAATTTAAGCAAAAGAAATAAGGAGAAAGAGAACTCAGAA	2858
DB	1679	AAGAAAGCTTATACTGAAATTTAAGCAAAAGAAATAAGGAGAAAGAGAACTCAGAA	1738
QY	2859	TTGGGGAGGCGATAGATTCCTATAAAATCAGCAAAATTTGTTGTAATTTAGAGGGAGAA	2918
DB	1739	TTGGGGAGGCGATAGATTCCTATAAAATCAGCAAAATTTGTTGTAATTTAGAGGGAGAA	1798
QY	2919	TTTGAATTAAGTATATAAAGGCGAATTTAGTATAGAGTACATTCATTAACATTTTGT	2978
DB	1799	TTTGAATTAAGTATATAAAGGCGAATTTAGTATAGAGTACATTCATTAACATTTTGT	1858
QY	2979	CAGATTTATTTCCGGTAAAAACCTGCTGAGCACTCTCATATACTAATTTAGTGTACATTTA	3038
DB	1859	CAGATTTATTTCCGGTAAAAACCTGCTGAGCACTCTCATATACTAATTTAGTGTACATTTA	1918
QY	3039	ACTTTGTATAATACAGAAATCTAAATATATTTAATGAATTCAGCAATATACACTTGACC	3098
DB	1919	ACTTTGTATAATACAGAAATCTAAATATATTTAATGAATTCAGCAATATACACTTGACC	1978
QY	3099	AGAAATTTGGAATTTCAAAATGTTGCTGCGGGTTATATACAGATGAGTACAGTAGTAG	3158
DB	1979	AGAAATTTGGAATTTCAAAATGTTGCTGCGGGTTATATACAGATGAGTACAGTAGTAG	2038
QY	3159	TTTATGTATACCAAGCTGGGTTATTTGCCAAGTTATATATACCAAAAGCTGTATGACTG	3218
DB	2039	TTTATGTATACCAAGCTGGGTTATTTGCCAAGTTATATATACCAAAAGCTGTATGACTG	2098
QY	3219	GATGTTCTGGTTACCTGGTTTACAAAATTTATCAGAGTAGT-----	3258
DB	2099	GATGTTCTGGTTACCTGGTTTACAAAATTTATCAGAGTAGTAAAACTTTGATATATGAG	2158
QY	3259	-----AAAACCTTTGATATATGAG	3278
DB	2159	GATATTTAAAACCTACATAGTATCATTTGATTCAGTAAACCTTTGATATATGAG	2218

Tue Nov 26 12:43:33 2002

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 162:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-09-030-607-162

Query Match 4.8% Score 165; DB 4; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.4e-36;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3159 TTTATGTATCACCAGACTGGGTTATTCGCAAGTTATATACCAAAAGCTGTATGACTG 3218
Db 165 TTTATGTATCACCAGACTGGGTTATTCGCAAGTTATATACCAAAAGCTGTATGACTG 106

QY 3219 GATGTTCTGGTTACCTGGTTTACAAAATATCAGAGTAGTAAACTTTGATATATATGAG 3278
Db 105 GATGTTCTGGTTACCTGGTTTACAAAATATCAGAGTAGTAAACTTTGATATATATGAG 46

QY 3279 GATATTAATAACTACACTAAGTAGTATCATTTGATTCGATTTCAGAAAGT 3323
Db 45 GATATTAATAACTACACTAAGTAGTATCATTTGATTCGATTTCAGAAAGT 1

RESULT 6
US-09-605-785-162/c
; Sequence 162, Application US/09605785
; Patent No. 6321716
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stoik, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C16
; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 835
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 162
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-605-785-162

QY 3159 TTTATGTATCACCAGACTGGGTTATTCGCAAGTTATATACCAAAAGCTGTATGACTG 3218
Db 165 TTTATGTATCACCAGACTGGGTTATTCGCAAGTTATATACCAAAAGCTGTATGACTG 106

QY 3219 GATGTTCTGGTTACCTGGTTTACAAAATATCAGAGTAGTAAACTTTGATATATATGAG 3278
Db 105 GATGTTCTGGTTACCTGGTTTACAAAATATCAGAGTAGTAAACTTTGATATATATGAG 46

QY 3279 GATATTAATAACTACACTAAGTAGTATCATTTGATTCGATTTCAGAAAGT 3323
Db 45 GATATTAATAACTACACTAAGTAGTATCATTTGATTCGATTTCAGAAAGT 1

RESULT 7
US-09-439-313-162/c
; Sequence 162, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 162
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-439-313-162

Query Match 4.8% Score 165; DB 4; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.4e-36;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3159 TTTATGTATCACCAGACTGGGTTATTCGCAAGTTATATACCAAAAGCTGTATGACTG 3218
Db 165 TTTATGTATCACCAGACTGGGTTATTCGCAAGTTATATACCAAAAGCTGTATGACTG 106

QY 3219 GATGTTCTGGTTACCTGGTTTACAAAATATCAGAGTAGTAAACTTTGATATATATGAG 3278
Db 105 GATGTTCTGGTTACCTGGTTTACAAAATATCAGAGTAGTAAACTTTGATATATATGAG 46

QY 3279 GATATTAATAACTACACTAAGTAGTATCATTTGATTCGATTTCAGAAAGT 3323
Db 45 GATATTAATAACTACACTAAGTAGTATCATTTGATTCGATTTCAGAAAGT 1

RESULT 8
US-09-352-616A-162/c
; Sequence 162, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yuqi
; APPLICANT: Xu, Jiangchun
; US-09-352-616A-162/c
```

```
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 162
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-352-616A-162

Query Match
Best Local Similarity 100.0%; Score 165; DB 4; Length 177;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3159 TTTATGTATCACCAGACTGGTATTGCGCAAGTTATATATACCAAAAGCTGTATGACTG 3218
Db 165 TTTATGTATCACCAGACTGGTATTGCGCAAGTTATATATACCAAAAGCTGTATGACTG 106

Qy 3219 GATGTTCTGGTTACCTGGTTTACAAATTTATCAGAGTAGTAAACTTTGATATATATGAG 3278
Db 105 GATGTTCTGGTTACCTGGTTTACAAATTTATCAGAGTAGTAAACTTTGATATATATGAG 46

Qy 3279 GATATTAAACTACACTAAGTATCATTTTCGATTCGATTCAGAAAGT 3323
Db 45 GATATTAAACTACACTAAGTATCATTTTCGATTCGATTCAGAAAGT 1

RESULT 9
US-09-232-149A-162/c
; Sequence 162, Application US/09232149A
; Patent No. 6465611
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C6
; CURRENT APPLICATION NUMBER: US/09/232,149A
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 162
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-232-149A-162

Query Match
Best Local Similarity 100.0%; Score 165; DB 4; Length 177;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3159 TTTATGTATCACCAGACTGGTATTGCGCAAGTTATATATACCAAAAGCTGTATGACTG 3218
Db 165 TTTATGTATCACCAGACTGGTATTGCGCAAGTTATATATACCAAAAGCTGTATGACTG 106

Qy 3219 GATGTTCTGGTTACCTGGTTTACAAATTTATCAGAGTAGTAAACTTTGATATATATGAG 3278
Db 105 GATGTTCTGGTTACCTGGTTTACAAATTTATCAGAGTAGTAAACTTTGATATATATGAG 46

Qy 3279 GATATTAAACTACACTAAGTATCATTTTCGATTCGATTCAGAAAGT 3323
Db 45 GATATTAAACTACACTAAGTATCATTTTCGATTCGATTCAGAAAGT 1

RESULT 10
US-08-990-571-64/c
; Sequence 64, Application US/08990571
; Patent No. 6214971
```

```
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G. et al.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B.
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,571
; FILING DATE: 11-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206)682-6031
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-990-571-64

Query Match
Best Local Similarity 1.7%; Score 59.8; DB 4; Length 285;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 405 ATCCATATACACCATGACCGACCATCCTACACACGAGCATCATCTGACCATCATGAG 464
Db 192 AACCATGTTCCACTAGGCCACCATCTCCTACTAGGCCACCATCTCCTACTAGGCCACCATG 133

Qy 465 CGTCACCTCAGACCATGACCATCCTCAGACCATCAGACCATCAGACCATCTGACCATCATCAC 524
Db 132 CTTCACTAGGCCACCATCTCCTACTAGGCCACCATCTCCTACTAGGCCACCATCTCCTAC 73

Qy 525 TCTCACCATATCAT 539
Db 72 TAGGCCACCATGCTT 58

RESULT 11
US-09-528-784A-64/c
; Sequence 64, Application US/09528784A
; Patent No. 6451315
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
; FILE REFERENCE: 210121.426C4
; CURRENT APPLICATION NUMBER: US/09/528,784A
; CURRENT FILING DATE: 10-03-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Babesia microti
```


Tue Nov 26 12:43:33 2002

; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-990-571-60

Query Match 1.7%; Score 59.8; DB 4; Length 351;
Best Local Similarity 65.2%; Pred. No. 5.4e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 405 ATCCATATACACCATGACCCAGCAGCATCTCAGACCCAGCAGCATCTCAGACCATGAG 464
DB 210 AACACAGCTTCACCTAGGCCACCCAGCTTCACCTAGGCCACCCAGCTTCACCTAGGCCACCCAG 151
QY 465 CGTCACTCAGACCATGAGCAGCATCTCAGACCCAGCAGCATCTCTGACCATGATCATCAC 524
DB 150 CTTCACTAGGCCACCCAGCTTCACCTAGGCCACCCAGCTTCACCTAGGCCACCCAGCTTCAC 91
QY 525 TCTCACCATAATCAT 539
DB 90 TAGGCCACCCAGCTT 76

RESULT 15

US-09-528-784A-60/C
; Sequence 60, Application US/09528784A
; Patent No. 6451315
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
; FILE REFERENCE: 210121.426C4
; CURRENT APPLICATION NUMBER: US/09/528,784A
; CURRENT FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Babesia microti
US-09-528-784A-60

Query Match 1.7%; Score 59.8; DB 4; Length 351;
Best Local Similarity 65.2%; Pred. No. 5.4e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 405 ATCCATATACACCATGACCCAGCAGCATCTCAGACCCAGCAGCATCTCAGACCATGAG 464
DB 210 AACACAGCTTCACCTAGGCCACCCAGCTTCACCTAGGCCACCCAGCTTCACCTAGGCCACCCAG 151
QY 465 CGTCACTCAGACCATGAGCAGCATCTCAGACCCAGCAGCATCTCTGACCATGATCATCAC 524
DB 150 CTTCACTAGGCCACCCAGCTTCACCTAGGCCACCCAGCTTCACCTAGGCCACCCAGCTTCAC 91
QY 525 TCTCACCATAATCAT 539
DB 90 TAGGCCACCCAGCTT 76

Search completed: November 8, 2002, 20:30:56
Job time : 109.498 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2002, 16:39:51 ; Search time 517.31 Seconds
(without alignments)
15066.735 Million cell updates/sec

Title: US-09-642-034-1
Perfect score: 3461
Sequence: 1 CTCGTGCCGAATCGGCAAG.....AAATTCCTATATCAGCTTG 3461

Scoring table: IDENTITY_NUC
Gapop 10.0 Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

Database :

N_Geneseq_101002.*

1: /SID2/cgcdata/geneseq/geneseq-emb1/NA1980.DAT.*

2: /SID2/cgcdata/geneseq/geneseq-emb1/NA1981.DAT.*

3: /SID2/cgcdata/geneseq/geneseq-emb1/NA1982.DAT.*

4: /SID2/cgcdata/geneseq/geneseq-emb1/NA1983.DAT.*

5: /SID2/cgcdata/geneseq/geneseq-emb1/NA1984.DAT.*

6: /SID2/cgcdata/geneseq/geneseq-emb1/NA1985.DAT.*

7: /SID2/cgcdata/geneseq/geneseq-emb1/NA1986.DAT.*

8: /SID2/cgcdata/geneseq/geneseq-emb1/NA1987.DAT.*

9: /SID2/cgcdata/geneseq/geneseq-emb1/NA1988.DAT.*

10: /SID2/cgcdata/geneseq/geneseq-emb1/NA1989.DAT.*

11: /SID2/cgcdata/geneseq/geneseq-emb1/NA1990.DAT.*

12: /SID2/cgcdata/geneseq/geneseq-emb1/NA1991.DAT.*

13: /SID2/cgcdata/geneseq/geneseq-emb1/NA1992.DAT.*

14: /SID2/cgcdata/geneseq/geneseq-emb1/NA1993.DAT.*

15: /SID2/cgcdata/geneseq/geneseq-emb1/NA1994.DAT.*

16: /SID2/cgcdata/geneseq/geneseq-emb1/NA1995.DAT.*

17: /SID2/cgcdata/geneseq/geneseq-emb1/NA1996.DAT.*

18: /SID2/cgcdata/geneseq/geneseq-emb1/NA1997.DAT.*

19: /SID2/cgcdata/geneseq/geneseq-emb1/NA1998.DAT.*

20: /SID2/cgcdata/geneseq/geneseq-emb1/NA1999.DAT.*

21: /SID2/cgcdata/geneseq/geneseq-emb1/NA2000.DAT.*

22: /SID2/cgcdata/geneseq/geneseq-emb1/NA2001A.DAT.*

23: /SID2/cgcdata/geneseq/geneseq-emb1/NA2001B.DAT.*

24: /SID2/cgcdata/geneseq/geneseq-emb1/NA2002.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3461	100.0	3461	24	ABK932205	Prostate cancer-as
2	3461	100.0	3461	24	ABA932299	Human breast cance
3	3391	98.0	3461	22	AAD13479	Human LIV-1 gene.
4	3335.6	96.4	3523	23	ABV25747	Human prostate exp
5	2761.2	79.8	2776	22	AAD133480	Human LIV-1-164847
6	2249	65.0	2404	19	AAT939071	Partial sequence o
7	2245.8	64.9	2404	17	AAT932010	Oestrogen-regulate
8	1486.8	43.0	1597	20	AAZ33622	Human breast tumou
9	1288	37.2	1310	19	AAT99070	Partial sequence o

10	1284.8	37.1	1310	17	AAT33219	Oestrogen-regulate
11	1141.4	33.0	1193	21	AAF16156	Human prostate can
12	1137.8	32.9	20778	22	AAK79819	Human immune/haem
13	619.6	17.9	899	24	ADA93300	Mouse breast cance
14	549.4	15.9	551	22	AAR13487	Human LIV-1-164647
15	404.8	11.7	483	24	ABL79043	Human ovarian can
c 16	390.8	11.3	395	22	AAH55569	Human breast tumou
c 17	390.4	11.3	424	22	AAH19686	Human breast cance
18	380.4	11.2	419	22	AAL22452	Human breast cance
19	364	10.5	366	21	AAC30456	Human secreted pro
20	360	10.4	385	21	AAA46009	Human metastatic m
21	340.2	9.8	603	23	ABV49440	Human prostate exp
22	319.2	9.2	337	24	ABK54054	Human head and nec
23	319	9.2	3309	23	ABV23278	Human prostate exp
24	319	9.2	3309	23	ABV29124	Human prostate exp
25	319	9.2	5332	22	ABNS9760	Novel human coding
26	319	9.2	5330	22	AAFP91861	Human secreted pro
27	267.6	7.7	537	23	ABV19673	Human prostate exp
c 28	262.4	7.6	423	23	AAI23559	Human breast cance
c 29	253.6	7.3	391	22	AAI13481	Probe to detect hu
c 30	251.4	7.3	469	22	AAU10106	Human breast cance
31	247	7.1	736	22	AAK92049	Human cDNA 5'-end
32	247	7.1	736	22	AAK93710	Human cDNA clone r
33	247	7.1	2863	22	AAK94381	Human full-length
c 34	240.8	7.0	413	22	AAI19430	Human breast cance
c 35	240.8	7.0	455	22	AAI09043	Human breast cance
c 36	240	6.9	241	22	AAAT47302	Human breast cance
c 37	240	6.9	241	22	AAF17872	Human breast cance
38	224.8	6.5	406	23	ABV60690	Human prostate exp
c 39	206.6	6.0	643	22	AAU18330	Human breast cance
c 40	189.4	5.5	405	22	AAU18130	Human breast cance
c 41	165	4.8	177	19	AAV61316	cDNA sequence of p
c 42	165	4.8	177	19	AAV58635	Prostate tumour sp
c 43	165	4.8	177	21	AAA06398	Human immunogeni
c 44	165	4.8	177	22	AAAG6306	Human prostate cDN
c 45	165	4.8	177	22	AA510157	Human prostate tum

ALIGNMENTS

RESULT 1	
ABK92205	
ID	ABK92205 standard; DNA; 3461 BP.
XX	
AC	ABK92205;
AC	
DT	15-AUG-2002 (first entry)
DT	
DE	Prostate cancer-associated DNA sequence #91.
DE	
OS	Prostate cancer; prostate tumour tissue; hum
KW	gene therapy; gene; ds.
KW	
XX	Mammalia.
OS	
PN	WC200230268-A2.
PN	
XX	18-APR-2002.
PD	
XX	12-OCT-2001; 2001WO-US32045.
PF	
XX	13-OCT-2000; 2000US-0687576.
XX	08-DEC-2000; 2000US-0733288.
PR	08-DEC-2000; 2000US-0733742.
PR	24-JAN-2001; 2001US-263957P.
PR	16-MAR-2001; 2001US-276291P.
PR	16-MAR-2001; 2001US-276888P.
PR	06-APR-2001; 2001US-281922P.
PR	24-APR-2001; 2001US-286214P.
PR	30-APR-2001; 2001US-0847046.
PR	04-MAY-2001; 2001US-288589P.
XX	
XX	

(EOSB-) EOS BIOTECHNOLOGY INC.

Glish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

WPI: 2002-471335/50.

P-PSDB: ABG61889;

Detecting a prostate cancer-associated transcript in a cell in a patient, useful for diagnosing prostate cancer (PC) or screening for PC, by determining if prostate cancer-associated genes are expressed in a prostate tissue -

Claim 22: page 374-375; 436pp; English.

The present invention relates to methods of detecting a prostate cancer-associated transcript in a cell from a patient. The method comprises contacting a biological sample from the patient with prostate cancer-associated polynucleotides (designated PC genes) that selectively hybridise to a sequence that is at least 80% identical to them. The prostate cancer-associated polynucleotide sequences are differentially expressed in prostate tumour tissue or in prostate cancer and are derived from the tissues of various organisms such as humans or other mammals (e.g. mice, sheep and dogs). The methods of the invention are useful for diagnosing and treating prostate cancer in mammals. The prostate cancer-associated genes are useful for diagnosing or treating prostate cancer, as well as for identifying modulators of prostate cancer or agents that inhibit prostate cancer. The nucleic acid sequences are particularly useful in gene therapy, as a vaccine or in antisense applications. ABK921115-ABK92263 represent prostate cancer-associated polynucleotide sequences.

Sequence 3461 BP: 1042 A: 686 C: 728 G: 1005 T: 0 other;

Query Match	100.0%;	Score 3461;	DB 24;	Length 3461;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3461:	Conservative	0;	Mismatches	0;
	Indels	0;	Gaps	0;

Qy	1	CTCGTGGCCGAATTCGGCAGACGACCGCTGTTCGCGCCCTGGTAGAGATTTCTCGAGACA	60
Db	1	CTCGTGGCCGAATTCGGCAGACGACCGCTGTTCGCGCCCTGGTAGAGATTTCTCGAGACA	60
Qy	61	CCAGTGGGCCCCGTGTGGAACAAACCTTGGCGCGTGGCCGGCCGTGGGACAAACAGGCC	120
Db	61	CCAGTGGGCCCCGTGTGGAACAAACCTTGGCGCGTGGCCGGCCGTGGGACAAACAGGCC	120
Qy	121	GCGGAGACGAAGGGCNAATGGCGAGAGAGTTATCTGTAATCTTGATCTCTGACCTTTGGCC	180
Db	121	GCGGAGACGAAGGGCNAATGGCGAGAGAGTTATCTGTAATCTTGATCTCTGACCTTTGGCC	180
Qy	181	TCTCTGTGCACAAATCCCCCTTCATGAACATAAAGCAGCTGCTTTCCGCCAGACCACTGAGA	240
Db	181	TCTCTGTGCACAAATCCCCCTTCATGAACATAAAGCAGCTGCTTTCCGCCAGACCACTGAGA	240
Qy	241	AAATTAGTCGGAATTCGGGAATCTGGCAATTAATGTTGACTTGGCAATTTCCACACGGCAAT	300
Db	241	AAATTAGTCGGAATTCGGGAATCTGGCAATTAATGTTGACTTGGCAATTTCCACACGGCAAT	300
Qy	301	ATCATCTACAACACGCTTTTCTACCGCTATGCGAGAAATAATCTTTGTGTCAGTTGAAGGT	360
Db	301	ATCATCTACAACACGCTTTTCTACCGCTATGCGAGAAATAATCTTTGTGTCAGTTGAAGGT	360
Qy	361	TCAGAAAATTACTTCAAAATATAGGCATAGATTAAGATTAAAGAAATCCATATACACCATG	420
Db	361	TCAGAAAATTACTTCAAAATATAGGCATAGATTAAGATTAAAGAAATCCATATACACCATG	420
Qy	421	ACCAGGACCATCTCAGACACGAGGCATCACTCAGACCATGAGCGCTCACTCAGACCATG	480
Db	421	ACCAGGACCATCTCAGACACGAGGCATCACTCAGACCATGAGCGCTCACTCAGACCATG	480
Qy	481	AGCATCACTCAGACACGAGCATCACTCTGACCATGATCATCACTCTCACCATAATCATG	540
Db	481	AGCATCACTCAGACACGAGCATCACTCTGACCATGATCATCACTCTCACCATAATCATG	540

Qy	1621	GAGCAGACTCACAGAGCCCTCCCAC	TTTGTATCTCAGACGCTTCAGTCTTTG	TTTGTATCTCAGACGCTTCAGTCTTTG	GAAGAAG	1681
Db	1621	GAGCAGACTCACAGAGCCCTCCCAC	TTTGTATCTCAGACGCTTCAGTCTTTG	TTTGTATCTCAGACGCTTCAGTCTTTG	GAAGAAG	1680
Qy	1681	AAGAGGTCATGATAGCTCATGCTCAT	CCACAGGAAGTCTACAATGAATATGT	TACCCACAG	1740	
Db	1681	AAGAGGTCATGATAGCTCATGCTCAT	CCACAGGAAGTCTACAATGAATATGT	TACCCACAG	1740	
Qy	1741	GGTGCAGAAATAATGCCATTACATATT	CCACAGATACACTCGGCCAGTCAGACG	ATCTCA	1800	
Db	1741	GGTGCAGAAATAATGCCATTACATATT	CCACAGATACACTCGGCCAGTCAGACG	ATCTCA	1800	
Qy	1801	TTCCACCACCATCATGACTACCATCAT	TATTTCTCCATCATCACACACCAAAACC	ACCATC	1860	
Db	1801	TTCCACCACCATCATGACTACCATCAT	TATTTCTCCATCATCACACACCAAAACC	ACCATC	1860	
Qy	1861	CTCACAGTCACAGCCGCTACTCTCGG	AGAGCTCAAGATGCCGGTCGCCACTT	1920		
Db	1861	CTCACAGTCACAGCCGCTACTCTCGG	AGAGCTCAAGATGCCGGTCGCCACTT	1920		
Qy	1921	TGGCCTGGATGGTGAATGGGTGATGG	CGCTGCACAAATTTTCAGCGATGGCCT	TAGCAATG	1980	
Db	1921	TGGCCTGGATGGTGAATGGGTGATGG	CGCTGCACAAATTTTCAGCGATGGCCT	TAGCAATG	1980	
Qy	1981	GTGCTGCTTTTTACTGAAGCGTTATCA	AGTGTTTAAAGTACTCTGTTGCTGTCT	GTCTGTC	2040	
Db	1981	GTGCTGCTTTTTACTGAAGCGTTATCA	AGTGTTTAAAGTACTCTGTTGCTGTCT	GTCTGTC	2040	
Qy	2041	ATGAGTTGCCCTCATGAATTAGTGACT	TTTGGCTTCTACTAAAGCTGGCATGACC	GGTTA	2100	
Db	2041	ATGAGTTGCCCTCATGAATTAGTGACT	TTTGGCTTCTACTAAAGCTGGCATGACC	GGTTA	2100	
Qy	2101	AGCAGGCTGCCCTTATATATGCAATGT	CAAGCCATGCTGGCGTATCTTGGAAAT	GCGCAACAG	2160	
Db	2101	AGCAGGCTGCCCTTATATATGCAATGT	CAAGCCATGCTGGCGTATCTTGGAAAT	GCGCAACAG	2160	
Qy	2161	GAATTTTTCATTTGGTCATATATGCT	GAAAAATTTTCTATGTGGATATTTGCA	CTTACTGCTG	2220	
Db	2161	GAATTTTTCATTTGGTCATATATGCT	GAAAAATTTTCTATGTGGATATTTGCA	CTTACTGCTG	2220	
Qy	2221	GCATTATCATGTATGTCTGGTTGATAT	GGTACCTGAAATGCTGCACAAATGATCT	A	2280	
Db	2221	GCATTATCATGTATGTCTGGTTGATAT	GGTACCTGAAATGCTGCACAAATGATCT	A	2280	
Qy	2281	GTGACCATGGATGTAGCCGCTGGGGAT	TATTTCTTTTACAGAAATGCTGGGATG	CTTTTGG	2340	
Db	2281	GTGACCATGGATGTAGCCGCTGGGGAT	TATTTCTTTTACAGAAATGCTGGGATG	CTTTTGG	2340	
Qy	2341	GTTTTTGGAAATTTGTTTACTATTATTT	CCATATTTTCAACATATAAATCGTGT	TTTCGTATAAAT	2400	
Db	2341	GTTTTTGGAAATTTGTTTACTATTATTT	CCATATTTTCAACATATAAATCGTGT	TTTCGTATAAAT	2400	
Qy	2401	TCATAGTTAAAGGTTTAAATGCTAGAT	GTCTAAAAGTTGTCAATGTTTCAGTAG	GGTCAT	2460	
Db	2401	TCATAGTTAAAGGTTTAAATGCTAGAT	GTCTAAAAGTTGTCAATGTTTCAGTAG	GGTCAT	2460	
Qy	2461	AGGGAGATGAGTTTGTATGCTGTACTAT	GCACGCGTTTAAAGTTAGTGGGTTTCTG	ATTT	2520	
Db	2461	AGGGAGATGAGTTTGTATGCTGTACTAT	GCACGCGTTTAAAGTTAGTGGGTTTCTG	ATTT	2520	
Qy	2521	TTGTATTTGAAATATTTGCTGTCTGTAT	TTCATAAGTCAGTTAAAGTTAATATTT	TAAGT	2580	
Db	2521	TTGTATTTGAAATATTTGCTGTCTGTAT	TTCATAAGTCAGTTAAAGTTAATATTT	TAAGT	2580	
Qy	2581	TATCTATCTTGGAGATAAATCTGTATGT	GCAATTCACCGGTATACCAGTTTATTA	TATG	2640	
Db	2581	TATCTATCTTGGAGATAAATCTGTATGT	GCAATTCACCGGTATACCAGTTTATTA	TATG	2640	
Qy	2641	TAAACAGAGATTTGGCATGACATGTTCT	GTATGTTTCAGGGAAAAATGCTTTTAAT	TGCT	2700	
Db	2641	TAAACAGAGATTTGGCATGACATGTTCT	GTATGTTTCAGGGAAAAATGCTTTTAAT	TGCT	2700	
Qy	2701	TTTTTCAAGAACTAACACAGTTATTTCT	TATPACTGGAATTTTATGCTCTCTCAAGA	ACTGCTG	2760	

D	b	2701	 TTTTTCAAGAACTAACACAGTTATTTCCTACTATCGGATTTTAGTGCTCTCAGAACA	TGCTGG	2760
Q	y	2761	 TGTTTAGGAATAAGATGTGCATGAAGCCATAAATACCAGAAAGCTTATACTGA	TTTA	2820
D	b	2761	 TGTTTAGGAATAAGATGTGCATGAAGCCATAAATACCAGAAAGCTTATACTGA	TTTA	2820
Q	y	2821	 AGCAAAGAAAATAAGGAGAAAGAGAGAAATCTGAGAAATTTGGGAGGCCATAGATCTTAT	TTAT	2880
D	b	2821	 AGCAAAGAAAATAAGGAGAAAGAGAGAAATCTGAGAAATTTGGGAGGCCATAGATCTTAT	TTAT	2880
Q	y	2881	 AAAAATCACAAAATTTGTTGTAATTTAGAGGGGAGAAAATTTAGAAATTAAGTATA	AAAAAGG	2940
D	b	2881	 AAAAATCACAAAATTTGTTGTAATTTAGAGGGGAGAAAATTTAGAAATTAAGTATA	AAAAAGG	2940
Q	y	2941	 CAGAAATTAGTAGAGTACATTTAAACATTTTGTGCAGGATATTTCCCGTAAAAAC	3000	
D	b	2941	 CAGAAATTAGTAGAGTACATTTAAACATTTTGTGCAGGATATTTCCCGTAAAAAC	3000	
Q	y	3001	 GTAGTGAGCACTCTCATATACATAATTTAGTGTACATTTAACTTTGTATAATACAGAAATCT	3060	
D	b	3001	 GTAGTGAGCACTCTCATATACATAATTTAGTGTACATTTAACTTTGTATAATACAGAAATCT	3060	
Q	y	3061	 AAATATATTTTAATGAATTCAGCAATATACACTTGACCAAGAAATTTGGAATTTCAAAATG	3120	
D	b	3061	 AAATATATTTTAATGAATTCAGCAATATACACTTGACCAAGAAATTTGGAATTTCAAAATG	3120	
Q	y	3121	 TTCTGCGGGTTATATACCAGATGAGTACAGTACAGTAGTATTATGTATCACCAGACTGGGT	3180	
D	b	3121	 TTCTGCGGGTTATATACCAGATGAGTACAGTACAGTAGTATTATGTATCACCAGACTGGGT	3180	
Q	y	3181	 TATTGCCAAGTTATATACCAAAAAGCTGTATGACTGGATCTCTGGTTACCTGGTTTA	3240	
D	b	3181	 TATTGCCAAGTTATATACCAAAAAGCTGTATGACTGGATCTCTGGTTACCTGGTTTA	3240	
Q	y	3241	 CAAAATTTATCAGAGTAGTAAACCTTTGATATATATACAGGATATTAATAACTACATAAGTA	3300	
D	b	3241	 CAAAATTTATCAGAGTAGTAAACCTTTGATATATATACAGGATATTAATAACTACATAAGTA	3300	
Q	y	3301	 TCATTTGATTCGATTACAGAAAGTACCTTTGATATCTCTCAGTGTCTCAGTGTATCATGT	3360	
D	b	3301	 TCATTTGATTCGATTACAGAAAGTACCTTTGATATCTCTCAGTGTCTCAGTGTATCATGT	3360	
Q	y	3361	 GAGCAATTTGCTTTATATACGGTACGTAGCCATACACTAGGCGCTGTCTGGCAATCTCTA	3420	
D	b	3361	 GAGCAATTTGCTTTATATACGGTACGTAGCCATACACTAGGCGCTGTCTGGCAATCTCTA	3420	
Q	y	3421	 GATCTTTCTTTTATACACAATAAATTCCTTATATCAGCTTG	3461	
D	b	3421	 GATCTTTCTTTTATACACAATAAATTCCTTATATCAGCTTG	3461	
RESULT 2					
ABA92299					
ID	XX	ABA92299	standard; cdna; 3461 BP.		
AC	XX	ABA92299;			
DT	XX	10-JUN-2002	(first entry)		
DE	XX	Human breast cancer 4 gene (BCR4) cdna.			
KW	XX	BCR4; human; breast cancer 4 gene; prostate cancer; diagnosis;			
KW	FT	gene therapy; vaccine; biochip; chromosome 18q12; gene; ss.			
OS	XX	Homo sapiens.			
FH	XX	Key	Location/Qualifiers		
FT	CD	CDS	138..2405		
ET	FT	/tag= a			
ET	FT	/product= "BCR4"			
FT	FT	138..197			
FT	FT	/*tag= b			

FT	mat_peptide	198..2402	
FT		/*tag= C	
XX			
PN	WO200216939-A2.		
PD	28-FEB-2002.		
PD			
PF	20-AUG-2001; 2001WO-US25997.		
PF			
PR	18-AUG-2000; 2000US-0642034.		
PR	08-DEC-2000; 2000US-0733320.		
XX			
XX	(EOSB-) EOS BIOTECHNOLOGY INC.		
PA			
PI	Mack D, Gish KC, Wilson KE;		
PI			
DR	WPI: 2002-242033/29.		
DR	P-PSDB; AAM51198.		
XX			
PT	Screening for drugs that affect expression of the breast cancer 4 gene		
PT	or its fragments, use of these to treat prostate and breast cancer, and		
PT	diagnosing these diseases -		
XX			
XX	Claim 27; Fig 1A-B; 83pp; English.		
PS			
CC	The presence sequence is that of cDNA identified for the human		
CC	breast cancer 4 gene (BCR4) on chromosome 18 (cytoband 18q12).		
CC	The gene was identified as being up-regulated in specific cancer		
CC	types in an oligonucleotide microarray interrogated with cRNAs		
CC	derived from multiple tissues. The gene encodes the protein given		
CC	in AAM51198. The invention provides claimed methods of using BCR4		
CC	polynucleotides, polypeptides and antibodies to: screen drug		
CC	candidates; screen bioactive agents capable of binding to BCR4;		
CC	evaluate the effect of a prostate cancer and/or breast cancer		
CC	drug; diagnose prostate cancer or breast cancer; screen for a		
CC	bioactive agent capable of interfering with the binding of BCR4 to		
CC	an antibody which binds BCR4; neutralise the effect of BCR4; treat		
CC	prostate cancer or breast cancer using an inhibitor of BCR4 (e.g.		
CC	an antibody); localise a therapeutic moiety to, or treat, prostate		
CC	cancer or breast cancer tissue by conjugating an antibody for BCR4		
CC	to a cytotoxic agent or radiolabel; inhibit prostate cancer or		
CC	breast cancer by administering an antisense molecule to the present		
CC	nucleic acid sequence; elicit an immune response by administering		
CC	BCR4 or a nucleic acid encoding BCR4; and determine the prognosis		
CC	of an individual with prostate cancer or breast cancer by		
CC	determining the level of BCR4 in a sample, a high level indicating		
CC	a poor prognosis. A biochip comprising BCR4 nucleic acids is also		
CC	claimed.		
XX			
SQ	Sequence 3461 BP; 1042 A; 686 C; 728 G; 1005 T; 0 other;		
	Query Match	100.0%;	Score 3461; DB 24; Length 3461;
	Best Local Similarity	100.0%;	Pred. No. 0;
	Matches 3461; Conservative	0;	Mismatches 0; Indels 0; Gaps
QY	1	CTCGTGGCGAAATTCGGCAGCAGACCGGGTGTTCGCCCGCTGTAGAGATTTCGGAAGACA	60
Db	1	CTCGTGGCGAAATTCGGCAGCAGACCGGGTGTTCGCCCGCTGTAGAGATTTCGGAAGACA	60
QY	61	CCAGTGGCGCCGTTGGGAACCAACCTTCGCGCGGTGGCGGCGGTGGGACCAACGAGGCC	120
Db	61	CCAGTGGCGCCGTTGGGAACCAACCTTCGCGCGGTGGCGGCGGTGGGACCAACGAGGCC	120
QY	121	GGCGAGAGCAAGGCGCAATGGCGAGGAAGTTATCTGTAAATCTTTGACTCTTGACCTTTGCC	180
Db	121	GGCGAGAGCAAGGCGCAATGGCGAGGAAGTTATCTGTAAATCTTTGACTCTTGACCTTTGCC	180
QY	181	TCTCTGTCAAAATCCCGCTTCATGAACCTAAAGCAGCTGCTTTCCCGCCAGACCACTTGAGA	240
Db	181	TCTCTGTCAAAATCCCGCTTCATGAACCTAAAGCAGCTGCTTTCCCGCCAGACCACTTGAGA	240
QY	241	AAATTTAGTCCGGAATTTGGGAATCTGGCAATTAATTTGACTTTGGCAATTTCCACACGCCAAT	300

QY 3417 TCTAGATGTTTCTTTTACACAATAAATTCCTTATATCAGC 3458
AA
Db 3404 TCTAGATGTTTCTTTTACACAATAAATTCCTTATATCAGC 3445

RESULT 5
AADI3480
ID AAD13480 standard; cDNA; 2776 BP.
XX AAD13480;

DT 06-NOV-2001 (first entry)

XX Human LIV-1-164647 cDNA.

XX Human; LIV-1-164647; cytostatic; estrogen-inducible gene; tumour;
KW cancer; breast; lung; prostate; colon; ovary; uterus; kidney; gastric;
XX salivary gland; carcinoma; drug screening; therapy; ss.
XX Homo sapiens.

XX Key Location/Qualifiers

PH CDS 73..2340

FT /*tag= a

FT /product= *Human LIV-1-164647 protein*

XX WO200155178-A2.

XX 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US02622.

XX 25-JAN-2000; 2000US-0177951.

PR 10-APR-2000; 2000US-0195761.

XX (GETH) GENENTECH INC.

XX Goddard A, Gurney AL, Smith V, Hongo JS, De Sauvage F;

PI WPI: 2001-502628/55.

DR P-PSDB; AAE07206.

XX New LIV-1 DNA and amino acids, useful for treating cancer or inhibiting
PT the proliferation of tumor cells in mammals, e.g. breast, lung,
PT prostate, colon, ovary, uterus, kidney, gastric or salivary carcinoma
PT .
XX Claim 6; Fig 2A; 150pp; English.

XX The present sequence is human estrogen-inducible LIV-1-164647

CC cDNA. LIV-1 is overexpressed in tumour tissues such as prostate,
CC colon, lung and breast. The LIV-1 DNA and polypeptide, and the monoclonal
CC antibody are useful for treating cancer and inhibiting the proliferation
CC of tumour cells in mammals, e.g. breast, lung, prostate, colon, ovary,
CC uterus, kidney, gastric or salivary carcinoma, or other tumour cell
CC types expressing the LIV-1-164647 protein. In particular, the mammal is
CC a human. The LIV-1 DNA and polypeptide may also be used in screening
CC assays for drug candidates.

XX Sequence 2776 BP; 823 A; 583 C; 592 G; 778 T; 0 other;

Query Match 79.8%; Score 2761.2; DB 22; Length 2776;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2766; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 68 GCCCGTGTGAACCAACCTGCGCGGTGGCCGGCGGTGGGACACAGGCGCGGGAGA 127

Db 3 GGCGGTGTGAACCAACCTGCGCGGTGGCCGGCGGTGGGACACAGGCGCGGGAGA 62

QY 128 CGAAGGCGCAATGCGGAGGAAGTTATCTGTAATCTTGATCCTGACCTTTGCCCTCTCTGT 187

Db 63 CGAAGGCGCAATGCGGAGGAAGTTATCTGTAATCTTGATCCTGACCTTTGCCCTCTCTGT 122

QY 188 CACAAATCCCCTTCTATGAACATAAAGCAGCTGCTTTCCCCCAGACACTGAGAAAATTAG 247
AA
Db 123 CACAAATCCCCTTCTATGAACATAAAGCAGCTGCTTTCCCCCAGACACTGAGAAAATTAG 182
QY 248 TCCGAATTGGGAATCTGGCATTAAATGTTGACTTGGCAATTTCCACAGGCAATATCATCT 307
AA
Db 183 TCCGAATTGGGAATCTGGCATTAAATGTTGACTTGGCAATTTCCACAGGCAATATCATCT 242
QY 308 ACACAGCTTTTCTACCGCTATGGAGAAAATAATTTCTTTGTGACTTGAAGGTTTCAGAAA 367
AA
Db 243 ACAACAGCTTTTCTACCGCTATGGAGAAAATAATTTCTTTGTGACTTGAAGGTTTCAGAAA 302
QY 368 ATTACTTCAAAATATAGGCATAGATTAAGATTAAAGAAATCCATATACACCATGACCAGA 427
AA
Db 303 ATTACTTCAAAATATAGGCATAGATTAAGATTAAAGAAATCCATATACACCATGACCAGA 362
QY 428 CCATCCTCAGACACGAGCATCTACGACCATGAGCGTCACTCAGACCATGAGCATCA 487
AA
Db 363 CCATCCTCAGACACGAGCATCTACGACCATGAGCGTCACTCAGACCATGAGCATCA 422
QY 488 CTCAGACACGAGCATCTCTGACCATGATCATCTCTCACCATATATCATGCTGCTTC 547
AA
Db 423 CTCAGACACGAGCATCTCTGACCATGATCATCTCTCACCATATATCATGCTGCTTC 482
QY 548 TGGTAAAAATTAAGCGAAAAGCTCTTTGCCAGACCATGACTCAGATAGTTCCAGGTAAGA 607
AA
Db 483 TGGTAAAAATTAAGCGAAAAGCTCTTTGCCAGACCATGACTCAGATAGTTCCAGGTAAGA 542
QY 608 TCCTAGAAACAGCCAGGGGAAAGAGCTCACCGACCAAGACATGCCAGTGTGTAGAGGAA 667
AA
Db 543 TCCTAGAAACAGCCAGGGGAAAGAGCTCACCGACCAAGACATGCCAGTGTGTAGAGGAA 602
QY 727 TGTCAAGGACAGTGTGTAGTGTAGTGAAGTGACCTCAACTGTGTACAACTGTCTCTGA 727
AA
Db 603 TGTCAAGGACAGTGTGTAGTGTAGTGAAGTGACCTCAACTGTGTACAACTGTCTCTGA 662
QY 728 AGGAATCACTTTCTAGAGACAATAGAGACTCCAAGACCTCGAAACCTTTCCCCCAAGA 787
AA
Db 663 AGGAATCACTTTCTAGAGACAATAGAGACTCCAAGACCTCGAAACCTTTCCCCCAAGA 722
QY 847 TGTAAAGCAGCTCCACTCCACCCAGTGTGCATCAAAAGCCGGGTGAGCCGGCTGGTGG 847
AA
Db 723 TGTAAAGCAGCTCCACTCCACCCAGTGTGCATCAAAAGCCGGGTGAGCCGGCTGGTGG 782
QY 907 TAGGAAAACAATGAATCTGTGAGTGAGCCCCCGAAAAGGCTTTTATGTTCCAGAAACAC 907
AA
Db 783 TAGGAAAACAATGAATCTGTGAGTGAGCCCCCGAAAAGGCTTTTATGTTCCAGAAACAC 842
QY 908 AAATGAAAATCCTCAGGAGTGTTCATGTCATCAAAAGCTACTGACATCTCATGCGATGGG 967
AA
Db 843 AAATGAAAATCCTCAGGAGTGTTCATGTCATCAAAAGCTACTGACATCTCATGGCATGGG 902
QY 1027 CATCCAGGTTCCGCTGGAATGAACAGAGTTCAACTATCTGTCCAGCCATCATCAACCA 1027
AA
Db 903 CATCCAGGTTCCGCTGGAATGAACAGAGTTCAACTATCTGTCCAGCCATCATCAACCA 962
QY 1087 AATTGATGCTAGATCTTCTGATTCATCAAGTGAAGGAGGCTGAAATCCCTCCAAA 1087
AA
Db 963 AATTGATGCTAGATCTTCTGATTCATCAAGTGAAGGAGGCTGAAATCCCTCCAAA 1022
QY 1088 GACCTATTTCATTAATAATAGCTGGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1147
AA
Db 1023 GACCTATTTCATTAATAATAGCTGGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1082
QY 1148 CTTGTCTCTGCTGGGGGTTATCTTTAGTGCCTCTCATGAATCGGGTGGTGGTGGTGGTGGTGG 1207
AA
Db 1083 CTTGTCTCTGCTGGGGGTTATCTTTAGTGCCTCTCATGAATCGGGTGGTGGTGGTGGTGGTGGTGG 1142
QY 1208 CTTGAGTTTCCTTGTGGGCACTGGCCGTTGGGCACTTTGAGTGGTGGTGGTGGTGGTGGTGGTGG 1267
AA
Db 1143 CTTGAGTTTCCTTGTGGGCACTGGCCGTTGGGCACTTTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1202
QY 1268 TCTTCCCAATTTCTCATGCAAGTCAACCATAGTATAGCCATAGTGAAGAACCAAGCAATGGA 1377
AA

1203 TCTTCCACATCTCATGCAAGTCCACCACATAGTCATAGGACGACCAATGGA 1262
1328 AATGAAAGAGGACCACTTTTCAGTCATCTGCTCTTCAAAACATAGAAGAAAGTGCCCTA 1387
1263 AATGAAAGAGGACCACTTTTCAGTCATCTGCTCTTCAAAACATAGAAGAAAGTGCCCTA 1322
1388 TTTTGTATTCACGCTGGAAGGGTCTAACAGCTCTAGGAGGCTGTATTTTCATGTTCTTGT 1447
1323 TTTTGTATTCACGCTGGAAGGGTCTAACAGCTCTAGGAGGCTGTATTTTCATGTTCTTGT 1382
1448 TGAACATGCTCCTACATTCATCAAAACATTTAAAGATAGAAGAAAGAAATCAGAAGAA 1507
1383 TGAACATGCTCCTACATTCATCAAAACATTTAAAGATAGAAGAAAGAAATCAGAAGAA 1442
1508 ACCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1567
1443 ACCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1502
1568 TTAACAAATGAGGAGAAAGTACATACATGATGATGATGATGATGATGATGATGATGATGATGAT 1627
1503 TTAACAAATGAGGAGAAAGTACATACATGATGATGATGATGATGATGATGATGATGATGATGAT 1562
1628 CTCAAGAGGCTCCACCTTTGATTTCTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1687
1563 CTCAGAGGCTCCACCTTTGATTTCTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1622
1688 CATGATGCTGCTCATCTCAGAGGAGTCTACATGATGATGATGATGATGATGATGATGATGATGAT 1747
1623 CATGATGCTGCTCATCTCAGAGGAGTCTACATGATGATGATGATGATGATGATGATGATGATGAT 1682
1748 GAATAAATGCCATTCACATTTCCACGATACATCGCGGAGTCAAGATGCTGCTGCTGCTGCTGCTGCT 1807
1683 GAATAAATGCCATTCACATTTCCACGATACATCGCGGAGTCAAGATGCTGCTGCTGCTGCTGCTGCT 1742
1808 CCATCATGACTACCATATTTCTCCATCATCACCACACCAACCAACCATCTCCACAG 1867
1743 CCATCATGACTACCATATTTCTCCATCATCACCACACCAACCAACCATCTCCACAG 1802
1868 TCACAGCAGGCTACTCTCGGAGGAGCTGAAAGATGCGGGGCTGCCACTTTGGGCTG 1927
1803 TCACAGCAGGCTACTCTCGGAGGAGCTGAAAGATGCGGGGCTGCCACTTTGGGCTG 1862
1928 GATGTGATTAATGGGTGATGGCTGCACAAATTCAGCGATGGCTAGCAATTTGGTGTGC 1987
1863 GATGTGATTAATGGGTGATGGCTGCACAAATTCAGCGATGGCTAGCAATTTGGTGTGC 1922
1988 TTTTACTGAAGCTTATCAAGTGGTTAAGTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2047
1923 TTTTACTGAAGCTTATCAAGTGGTTAAGTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1982
2048 GCCTCATGAATAGGTGACTTTGCTGTTCTACTAAGCTGCGATGACCGTTAAGCAGGC 2107
1983 GCCTCATGAATAGGTGACTTTGCTGTTCTACTAAGCTGCGATGACCGTTAAGCAGGC 2042
2108 TGTCTTTTAAATGCAATGTCAGCCATGCTGCGGTATCTTGGAAATGCGCAACAGGAATTTT 2167
2043 TGTCTTTTAAATGCAATGTCAGCCATGCTGCGGTATCTTGGAAATGCGCAACAGGAATTTT 2102
2168 CATTTGCTCATTTGCTGAAATGTTTCTATGCGATATTTGCACTTACTGCTGGCTTAT 2227
2103 CATTTGCTCATTTGCTGAAATGTTTCTATGCGATATTTGCACTTACTGCTGGCTTAT 2162
2228 CATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2287
2163 CATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2222
2288 TGGATGATAGCGCTGGGGGTATTTCTTTTACAGATGCTGGGATGCTTTGGGTTTGG 2347
2223 TGGATGATAGCGCTGGGGGTATTTCTTTTACAGATGCTGGGATGCTTTGGGTTTGG 2282
2348 AATTATGTTACTTATTTCCATATTTTGAACATATAAATCGTGTTCGTATATAATTTCTAGTT 2407

2283 AATTATGTTACTTATTTCCATATTTTGAACATATAAATCGTGTTCGTATATAATTTCTAGTT 2342
2408 AAGGTTTAAATGCTAGAGTAGCTTTAAAGAGTTGTGCATAGTTTCCAGTAGGTCATAGGAGA 2467
2343 AAGGTTTAAATGCTAGAGTAGCTTTAAAGAGTTGTGCATAGTTTCCAGTAGGTCATAGGAGA 2402
2468 TGAGTTTGTATGCTGCTACTATGCGAGCGTTTAAAGCTTAGTGGGTTTCTGATTTTCTGATT 2527
2403 TGAGTTTGTATGCTGCTACTATGCGAGCGTTTAAAGCTTAGTGGGTTTCTGATTTTCTGATT 2462
2528 GAATATGCTGCTGCTTACAAAGTCAGTTAAAGGTACGTTTAAATATATTTAAGTTATTCTA 2587
2463 GAATATGCTGCTGCTTACAAAGTCAGTTAAAGGTACGTTTAAATATATTTAAGTTATTCTA 2522
2588 TCTTGAGATATAATGCTGCTATGTCGAATTCACCGGTATTACCAAGTTATTATGTAACAA 2647
2523 TCTTGAGATATAATGCTGCTATGTCGAATTCACCGGTATTACCAAGTTATTATGTAACAA 2582
2648 GAGATTTGGCATGCATGTTCTGATGTTTTCAGGGAAAAATGCTTTAATGCTTTTCA 2707
2583 GAGATTTGGCATGCATGTTCTGATGTTTTCAGGGAAAAATGCTTTAATGCTTTTCA 2642
2708 GAACTAACACAGTTATTCCTATCTGATTTAGGTTTCTGGAAGTCTGCTGGTGTAG 2767
2643 GAACTAACACAGTTATTCCTATCTGATTTAGGTTTCTGGAAGTCTGCTGGTGTAG 2702
2768 GAATAAGATGTCATGAAGCCTTAAATACCAAGAAAGCTTATCTGAAATTTAAGCAAG 2827
2703 GAATAAGATGTCATGAAGCCTTAAATACCAAGAAAGCTTATCTGAAATTTAAGCAAG 2762
2828 AATAAAGGAGAA 2841
2763 AAAAAAAAAAAAA 2776
RESULT 6
AAT99071 standard; cDNA; 2404 BP.
XX AAT99071;
XX 24-MAR-1998 (first entry)
XX Partial sequence of the pLIV1 gene.
XX Oestrogen regulated gene; pLIV1 gene partial sequence; breast cancer;
KW metastatic risk assessment; hormone therapy responsiveness;
KW tumour oestrogen receptor expression; lymph node involvement; ds.
OS Homo sapiens.
XX Key Location/Qualifiers
FH 1..1296
FT /*tag= a
XX US5693465-A.
XX 02-DEC-1997.
XX 22-SEP-1994; 94US-0311023.
XX 22-SEP-1994; 94US-0311023.
XX (UYWA-) UNIV WALES COLLEGE OF MEDICINE.
XX Gee JMW, Green CD, Manning DL, Nicholson RI;
XX WPI; 1998-031740/03.
XX P-PSDB; AAW34528.
XX Assessment of metastatic risk or oestrogen responsive-ness in breast
XX cancer - by detecting expression of pLIV1 protein

PS Disclosure: Column 11-14; 1lpp; English.

xx This sequence represents a partial sequence of the pLIV1 gene. This sequence extends further in the 3' direction than the pLIV1 partial sequence shown in A099070. The protein encoded by this sequence is an oestrogen regulated gene associated with breast cancer. The protein encoded by this sequence can be used in the method of the invention. The method is for determining the risk of metastasis of a female breast tumour and predicting the responsiveness of a female breast tumour to hormone treatment comprise determining if a tissue sample from the tumour expresses a polypeptide comprising at least 14 continuous amino acids of the protein encoded by this sequence. The method is useful for determining the risk of metastasis of a female breast tumour and to predict the response of the female breast tumour to hormone treatment. pLIV1 expression is highly predictive of tumour oestrogen receptor expression and lymph node involvement.

xx

SQ Sequence 2404 BP; 726 A; 411 C; 492 G; 775 T; 0 other;

Query Match		65.0%; Score 2249; DB 19; Length 2404;
Best Local Similarity		97.4%; Pred. No. 0;
Matches 2341; Conservative		0; Mismatches 0; Indels 62; Gaps 3;
QY	1119	GGTTTATAGCCATTTCATCATCAGTTTCTGCTGCTGGGGTTATCTTAGTGCCT 1178
DB	1	GGTTTATAGCCATTTCATCATCAGTTTCTGCTGCTGGGGTTATCTTAGTGCCT 60
QY	1179	CTCATGAATCGGCTGTTTTTCAAAATTCCTGAGTTTCTGTCGACCTGGCGTTGGG 1238
DB	61	CTCATGAATCGGCTGTTTTTCAAAATTCCTGAGTTTCTGTCGACCTGGCGTTGGG 120
QY	1239	ACTTTGAGTGTGATGCTTTTTTACACCTTTTCCACATCTCTCATGCAAGTCACCCACAT 1298
DB	121	ACTTTGAGTGTGATGCTTTTTTACACCTTTTCCACATCTCTCATGCAAGTCACCCACAT 180
QY	1299	AGTCATAGCCATGAAGAACAGCAATGGAATGAAATGAAAGAGGACCCATTTTCAGTCATCTG 1358
DB	181	AGTCATAGCCATGAAGAACAGCAATGGAATGAAATGAAAGAGGACCCATTTTCAGTCATCTG 240
QY	1359	TCCTTCTCAAAACATAGAGAAAGTGCCTATTTTCATTTCCACGTGGAAGGGTCTAACAGCT 1418
DB	241	TCCTTCTCAAAACATAGAGAAAGTGCCTATTTTCATTTCCACGTGGAAGGGTCTAACAGCT 300
QY	1419	CTAGAGGCGCTGATTTTCATGTTTCTTGTGAAACATGCTCTCACATTTGATCAACAATTT 1478
DB	301	CTAGAGGCGCTGATTTTCATGTTTCTTGTGAAACATGCTCTCACATTTGATCAACAATTT 360
QY	1479	AAAGATAGAGAAAGAAAGATCAGAGAAACCTGAAATGATGATGATGTTGGAGATTAG 1538
DB	361	AAAGATAGAGAAAGAAAGATCAGAGAAACCTGAAATGATGATGATGTTGGAGATTAG 420
QY	1539	AAGCAGTTGTCCAAAGTATGAATCTCAACTTTTCAACAAATGAGGAGAAAGTAGATACAGAT 1598
DB	421	AAGCAGTTGTCCAAAGTATGAATCTCAACTTTTCAACAAATGAGGAGAAAGTAGATACAGAT 480
QY	1599	GATCGAAGTGAAGGCTATTTACGAGAGAGACTCACAAGAGCCCTCCCATTTGATCTCAG 1658
DB	481	GATCGAAGTGAAGGCTATTTACGAGAGAGACTCACAAGAGCCCTCCCATTTGATCTCAG 540
QY	1659	CAGCTCGAGCTTGGAGAGAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATG 1718
DB	541	CAGCTCGAGCTTGGAGAGAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY	1719	TACAAATGAATATGTACCCAGAGGGTGCAGAAATAAATGCCATTTACATTTCCACGATACA 1778
DB	601	TACAAATGAATATGTACCCAGAGGGTGCAGAAATAAATGCCATTTACATTTCCACGATACA 660
QY	1779	CTCGGCCAGTCAGAGGATCTCATTTACACCATCATGACTACCATCATATTTCCATCAT 1838
DB	661	CTCGGCCAGTCAGAGGATCTCATTTACACCATCATGACTACCATCATATTTCCATCAT 720
QY	1839	CACCACCAACCAACCACTCTTCACAGTCACAGCCGCTACTCTCGGGAGGAGCTG 1898
DB		

DB	721	CACCACCAACCAACCAACCATCTCAGAGTCACAGCAGCGCTACTCTCGGGAGGAGCTG 780
QY	1899	AAAGATGCGCGCGTGCACCATTTGGCCCTGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1958
DB	781	AAAGATGCGCGCGTGCACCATTTGGCCCTGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 840
QY	1959	TTCAAGCGATGGCGCTAGCAATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2018
DB	841	TTCAAGCGATGGCGCTAGCAATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
QY	2019	ACTTCTGTTGCTGTTGTTCTGTCATGAGTTGCTCATGAATTAGTGACTTTGCTGCTTCTTA 2078
DB	901	ACTTCTGTTGCTGTTGTTCTGTCATGAGTTGCTCATGAATTAGTGACTTTGCTGCTTCTTA 960
QY	2079	CTAAAGGCTGGCATGACCGTTAAGCAGGCTGCTTCTTATTAATGCAATTTGTCAGCCCATGCTG 2138
DB	961	CTAAAGGCTGGCATGACCGTTAAGCAGGCTGCTTCTTATTAATGCAATTTGTCAGCCCATGCTG 1020
QY	2139	GGGTATCTTTGGAATGGCAACAGCAATTTTCATTTGGTGCATTTATGCTGAAATGTTTCTATG 2198
DB	1021	GGGTATCTTTGGAATGGCAACAGCAATTTTCATTTGGTGCATTTATGCTGAAATGTTTCTATG 1080
QY	2199	TGATATTTGCACTTACTGCTGGCTTATTCATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2258
DB	1081	TGATATTTGCACTTACTGCTGGCTTATTCATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY	2259	GAAATGCTGCACAAATGATGCTAGTGACCATGATGCTAGTGACCATGATGCTAGTGACCATGATG 2318
DB	1141	GAAATGCTGCACAAATGATGCTAGTGACCATGATGCTAGTGACCATGATGCTAGTGACCATGATG 1200
QY	2319	CAGATGCTGGGATGCTTTTGGGTTTGGGAATATGTTACTTATTTCCATATTTGAACAT 2378
DB	1201	CAGATGCTGGGATGCTTTTGGGTTTGGGAATATGTTACTTATTTCCATATTTGAACAT 1259
QY	2379	AAATGCTGTTTCTGATATAAATTTTCTAGTTTAAAGTTTAAATGCTAGAGTACGTTTAAAAAGT 2438
DB	1260	AAATGCTGTTTCTGATATAAATTTTCTAGTTTAAAGTTTAAATGCTAGAGTACGTTTAAAAAGT 1318
QY	2439	TGTCATAGTTTTCAGTAGGTCATAGGAGATGATGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2498
DB	1319	TGTCATAGTTTTCAGTAGGTCATAGGAGATGATGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1378
QY	2499	AAATGCTGGGTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGAT 2558
DB	1379	AAATGCTGGGTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGAT 1438
QY	2559	AGTACGTTTAAATATTTAAAGTTTCTATCTTGAGATATAAATCTGATGCTGCAATTTCA 2618
DB	1439	AGTACGTTTAAATATTTAAAGTTTCTATCTTGAGATATAAATCTGATGCTGCAATTTCA 1498
QY	2619	CCGCTATTACCACTTTTATTTATGTAACAAGATTTGGCATGACATGTTCTGCTGATGTTTCT 2678
DB	1499	CCGCTATTACCACTTTTATTTATGTAACAAGATTTGGCATGACATGTTCTGCTGATGTTTCT 1558
QY	2679	AGGAAAAATGCTTTTAAATGCTTTTCAAGAACTAACACAGATTTTCTTATCTACTGGATTT 2738
DB	1559	AGGAAAAATGCTTTTAAATGCTTTTCAAGAACTAACACAGATTTTCTTATCTACTGGATTT 1618
QY	2739	TAGGCTCTCTCAAGCACTGCTGCTGTTTAAAGAACTAACACAGATTTTCTTATCTACTGGATTT 2798
DB	1619	TAGGCTCTCTCAAGCACTGCTGCTGTTTAAAGAACTAACACAGATTTTCTTATCTACTGGATTT 1678
QY	2799	AGAAAACTTATCTGAATTTAAGCAACAAATTAAGGAGAGAAAGAGAGAAATCTGAGAA 2858
DB	1679	AGAAAACTTATCTGAATTTAAGCAACAAATTAAGGAGAGAGAAAGAGAGAAATCTGAGAA 1738
QY	2859	TTGGGAGGATAGATTTCTTATAAAAATCAGAAAAATTTGCTGTAATTTAGAGGGGAGAAA 2918
DB	1739	TTGGGAGGATAGATTTCTTATAAAAATCAGAAAAATTTGCTGTAATTTAGAGGGGAGAAA 1798
QY	2919	TTTAGAATTTAAGTATAAAAAGGCAATTTAGTATAGAGTACATTTCAATTAACATTTTGT 2978
DB	1799	TTTAGAATTTAAGTATAAAAAGGCAATTTAGTATAGAGTACATTTCAATTAACATTTTGT 1858

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E, Rosentahl A;
PI WPI; 1999-528981/45.
XXXX Human nucleic acid sequences and protein products from tumor breast
PT tissue, useful for breast cancer therapy -
XX

XX Claim 3; 93; 188pp; German.

XX This invention describes novel human nucleic acid sequences from tumor
CC breast tissue which have cytosolic activity. The nucleic acid sequences
CC can be used to produce and isolate full-length gene sequences. They can
CC be used to express proteins, which can be used as tools to find an
CC activity against breast cancer. The sequences can be used in sense or
CC antisense form. They are especially useful for medicaments for gene
CC therapy to treat breast cancer. AA233611-248617 represents expressed
CC sequence tags described in the method of the invention.

XX Sequence 1597 BP; 476 A; 229 C; 337 G; 555 T; 0 other;

Query Match 43.0%; Score 1486.8; DB 20; Length 1597;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1559; Conservative 0; Mismatches 7; Indels 8; Gaps 6;

QY 1895 GCTGAAGATGCGGGCTGCCACTTTGGCTGGATGGTGAATGGGTGATGGCGCTGCA 1954

DB 19 GCTGAAGATGCGGGCTGCCACTTTGGCTGGATGGTGAATGGGTGATGGCGCTGCA 78

QY 1955 CAATTTGAGGATGGCTAGCAATTTGGTCTGCTTTTACTGAGGCTTATCAAGTGGTTT 2014

DB 79 CAATTTGAGGATGGCTAGCAATTTGGTCTGCTTTTACTGAGGCTTATCAAGTGGTTT 138

QY 2015 AAGTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2074

DB 139 AAGTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 198

QY 2075 TCTACTAAAGGCTGGATGACCTTAAAGCAGGCTGCTTTTAAATGCAATGTGACCCAT 2134

DB 199 TCTACTAAAGGCTGGATGACCTTAAAGCAGGCTGCTTTTAAATGCAATGTGACCCAT 258

QY 2135 GCTGGGCTATCTTGGGATGCAACAGGAATTTTCATTGGTTCATTATGCTGAAATGTTTC 2194

DB 259 GCTGGGCTATCTTGGGATGCAACAGGAATTTTCATTGGTTCATTATGCTGAAATGTTTC 318

QY 2195 TATGTGATATTTGCACTACTGCTGGCTTATTCATGTATGTGCTGCTGCTGCTGCTGCT 2254

DB 319 TATGTGATATTTGCACTACTGCTGGCTTATTCATGTATGTGCTGCTGCTGCTGCTGCT 378

QY 2255 ACCTGAAATGCTGCACATGATGCTAGTGACCATGGATGACCGCTGGGGTATTTCTTT 2314

DB 379 ACCTGAAATGCTGCACATGATGCTAGTGACCATGGATGACCGCTGGGGTATTTCTTT 438

QY 2315 TTTACAGAAATGCTGGGATGCTTTGGGTTTGGAAATATGTTACTTATTTCCATATTTCA 2374

DB 439 TTTACAGAAATGCTGGGATGCTTTGGGTTTGGAAATATGTTACTTATTTCCATATTTCA 498

QY 2375 ACATAAAATGCTGTTTCGTATATAAATTTCTAGTTAAGGTTTAAATGCTAGAGTAGCTTAA 2434

DB 499 ACATAAAATGCTGTTTCGTATATAAATTTCTAGTTAAGGTTTAAATGCTAGAGTAGCTTAA 558

QY 2435 AAGTTCTCATAGTTTTCAGTAGGTCATAGGAGATGATTTTCTATGCTGATGACGG 2494

DB 559 AAGTTCTCATAGTTTTCAGTAGGTCATAGGAGATGATTTTCTATGCTGATGACGG 618

QY 2495 TTTAAAGTTAGTGGGTTTGTGATTTTGTGATATTTGCTGCTGCTGCTGCTGCTGCTGCT 2554

DB 619 TTTAAAGTTAGTGGGTTTGTGATTTTGTGATATTTGCTGCTGCTGCTGCTGCTGCTGCT 678

QY 2555 TTTAAAGTTAGTGGTATTTTAAATTTTAAATTTTCTGAGATAAATCTGATGTGCA 2614

DB 679 TTTAAAGTTAGTGGTATTTTAAATTTTAAATTTTCTGAGATAAATCTGATGTGCA 738

QY 2615 TTCACCGGTATTACCAGTTTATTATGTAACAAGAGATTGGCATGACATGTTCTGTATG 2674
DB 739 TTCACCGGTATTACCAGTTTATTATGTAACAAGAGATTGGCATGACATGTTCTGTATG 798
QY 2675 TTTACAGGAAAAAATGCTTTTAAATGCTTTTTCAGAACTAAACACAGTTTCTCTATACCTG 2734
DB 799 TTTACAGGAAAAAATGCTTTTAAATGCTTTTTCAGAACTAAACACAGTTTCTCTATACCTG 858
QY 2735 ATTTTAGTCTCTGAAGAACTGCTGCTGTTTAGGAATTAAGAAATGTCATGAAGCCTAAAA 2794
DB 859 ATTTTAGTCTCTGAAGAACTGCTGCTGTTTAGGAATTAAGAAATGTCATGAAGCCTAAAA 918
QY 2795 TACCAGAAAGCTTATACCTGAATTTAAGCAAGAAATAAGGAGAAAGAGAGAACTG 2854
DB 919 TACCAGAAAGCTTATACCTGAATTTAAGCAAGAAATAAGGAGAAAGAGAGAACTG 978
QY 2855 AGAATTTGGGAGGCTATAGATCTTTATAAATAACAAAAATTTGTTGTAATTTAGAGGGA 2914
DB 979 AGAATTTGGGAGGCTATAGATCTTTATAAATAACAAAAATTTGTTGTAATTTAGAGGGA 1038
QY 2915 GAAATTTAGAAATTAAGTATAAAGGCGCAATTTAGTATAGAGTACATTCATTTAAACATTT 2974
DB 1039 GAAATTTAGAAATTAAGTATAAAGGCGCAATTTAGTATAGAGTACATTCATTTAAACATTT 1098
QY 2975 TTGTGAGGATTTATCCCGTAAAAACGTAGTGAG - CACTCTCATATATACTAATT - AGTG 3030
DB 1099 TTGTGAGGATTTATCCCGTAAAAACGTAGTGAGGCACTTTTCATATACTAATTTAGTTG 1158
QY 3031 TACATTTAACTTTGTATTAATACAGAAATCTAAATATATTTAATGAATTTCAAGCAATATA - 3089
DB 1159 TACATTTAACTTTGTATTAATACAGAAATCTAAATATATTTAATGAATTTCAAGCAATATA 1218
QY 3090 CACTTGACCAAGAAATTTGAAATTTTCAAAATGTTCTGCGGGTTATATACAGATGAGTAC 3149
DB 1219 CACTTGACCAAGAAATTTGAAATTTTCAAAATGTTCTGCGGG - TATATACAGATGAGTAC 1277
QY 3150 AGTGAGTAG - TTTATGTATCACAGACTGGGTTATGCGCAAGTTATATATACCAAAAGC 3208
DB 1278 AGTGAGTAGTTTATGTATCACAGACTGGGTTATGCGCAAGTTATATATACCAAAAGC 1337
QY 3209 TGTATGACTGGATGTTCTGTTTACCTGTTTACAAAAATATACAGAGTAGTAAAACTTTGA 3268
DB 1338 TGTATGACTGGATGTTCTGTTTACCTGTTTACAAAAATATACAGAGTAGTAAAACTTTGA 1397
QY 3269 TATATATCAGGATATTAACACTTACACTAAGTATCATCTTTCGATTCAGAAAGTACTTTT 3328
DB 1398 TATATATCAGGATATTAACACTTACACTAAGTATCATCTTTCGATTCAGAAAGTACTTTT 1457
QY 3329 GATATCTCTCAGTCTCAGTCTATCATCTGAGCAATTTGTC - TTTATATACGCTACTG 3387
DB 1458 GATATCTCTCAGTCTCAGTCTATCATCTGAGCAATTTGTCATATATACGCTACTG 1517
QY 3388 TAGCCATACCTAGGCTGCTGCTGGCATCTCTAGATGTTCTTTTACACAATAAATTC 3447
DB 1518 TAGCCATACCTAGGCTGCTGCTGGCATCTCTAGATGTTCTTTTACACAATAAATTC 1577
QY 3448 CTTATATCAGCTTG 3461
DB 1578 CTTATATCAGCTTG 1591

RESULT 9

AAT99070

ID AAT99070 standard; cDNA; 1310 BP.

XX AAT99070;

AC AAT99070;

XX 24-MAR-1998 (first entry)

XX Partial sequence of the pLIV1 gene.

XX Oestrogen regulated gene; pLIV1 gene partial sequence; breast cancer;

KW

XX

XX

DE	Human prostate cancer antigen nucleotide sequence SEQ ID NO:591.	
XX	Human; prostate cancer; prostate cancer antigen; detection; diagnosis;	
KW	neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;	
KW	vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;	
KW	antibacterial; gene therapy; neural; immune; reproductive; renal;	
KW	gastrointestinal; pulmonary; cardiovascular; proliferative disorder;	
KW	wound; infectious disease; ss.	
OS	Homo sapiens.	
XX	WO200055174-A1.	
PN	21-SEP-2000.	
XX	08-MAR-2000; 2000WO-US05988.	
XX	12-MAR-1999; 99US-0124270.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	(ROSE/) ROSEN C A.	
XX	Rosen CA, Ruben SM;	
PI	WPI; 2000-587513/55.	
XX	P-PSDB; AAB56953.	
DR	Prostate cancer associated gene sequences, referred to as prostate	
DR	cancer antigens, useful for treatment, prevention, and diagnosis of	
PT	disorders such as prostate cancer	
PT		
XX	Claim 1; Page 1045-1046; 2338pp; English.	
PS		
XX	AAF15566 to AAF16505 encode the human prostate cancer associated	
CC	proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.	
CC	The prostate cancer antigens can have neuroprotective, cytostatic,	
CC	cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,	
CC	nephrotropic, antiinfective, gynaecological and antibacterial activities,	
CC	and can be used in gene therapy. The prostate cancer antigen	
CC	polynucleotides may be used for detection of prostate cancer, chromosome	
CC	identification, as chromosome markers, and for numerous other diagnostic	
CC	or research purposes. The prostate cancer antigens may be used to treat	
CC	disorders such as neural, immune, muscular, reproductive,	
CC	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative	
CC	disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to	
CC	AAF57303 represent sequences used in the exemplification of the present	
CC	invention.	
XX		
SQ	Sequence 1193 BP; 393 A; 259 C; 235 G; 302 T; 4 other;	
	Query Match 33.0%; Score 1141.4; DB 21; Length 1193;	
	Best Local Similarity 99.7%; Pred. No. 5.2e-274;	
	Matches 1139; Conservative 4; Mismatches 0; Indels 0; Gaps 0;	
QY	676 ACAGTGTAGTCTAGTGAAGTGACCTCAACTGTGTACAACTGTCTCTGAAGGAATC 735	
DB	1 ACAGTGTAGTCTAGTGAAGTGACCTCAACTGTGTACAACTGTCTCTGAAGGAATC 60	
QY	736 ACTTCTTAGAGCAATAGAGACTCCAGACCTGGAAACTCTTCCCAAGATGTAAACA 795	
DB	61 ACTTCTTAGAGCAATAGAGACTCCAGACCTGGAAACTCTTCCCAAGATGTAAACA 120	
QY	796 GCTCCACTCCACCAGGTGTACATCAAGAGCGGGTGTAGCGGGTGTAGGAAATA 855	
DB	121 GCTCCACTCCACCAGGTGTACATCAAGAGCGGGTGTAGCGGGTGTAGGAAATA 180	
QY	856 CAAATGAATCTGTGAGTGTAGCGGGTGTAGCGGGTGTAGCGGGTGTAGGAAATA 915	
DB	181 CAAATGAATCTGTGAGTGTAGCGGGTGTAGCGGGTGTAGCGGGTGTAGGAAATA 240	
QY	916 ATCCCTCAGGAGTGTTCATGATGATCAAAAGCTACTGACATCTCATGGCATGGGCATCCAGG 975	
DB	241 ATCCCTCAGGAGTGTTCATGATGATCAAAAGCTACTGACATCTCATGGCATGGGCATCCAGG 300	

QY	976 TTCCGCTGAATGCAACAGAGTTCAACTATCTCTGTCCAGCCATCATCAACCAAAATGATG 1035	
DB	301 TTCCGCTGAATGCAACAGAGTTCAACTATCTCTGTCCAGCCATCATCAACCAAAATGATG 360	
QY	1036 CTAGATCTTGTCTGATTCATACAAAGTGAAGAGCGCTGAATCCCTCCAAAGACCTATT 1095	
DB	361 CTAGATCTTGTCTGATTCATACAAAGTGAAGAGCGCTGAATCCCTCCAAAGACCTATT 420	
QY	1096 CATTACAAATAGCCTGGGTTGGTGGTGTATATAGCCATTTCCATCATCATGTTCCCTGCTC 1155	
DB	421 CATTACAAATAGCCTGGTGGTGTATATAGCCATTTCCATCATCATGTTCCCTGCTC 480	
QY	1156 TGTGGGGGTTATCTTGTGCTCTCATGAATCGGGTGTCTTCAAAATTCCTCCGAGCTT 1215	
DB	481 TGTGGGGGTTATCTTGTGCTCTCATGAATCGGGTGTCTTCAAAATTCCTCCGAGCTT 540	
QY	1216 TCTTGTGGCACTGGCCGTTGGGACTTTTCAGTGGTGTATGCTTTTTCACACCTTCTTCCAC 1275	
DB	541 YCCYGTGGCACTGGCCGTTGGGACTTTTCAGTGGTGTATGCTTTTTCACACCTTCTTCCAC 600	
QY	1276 ATTCTCATGCAAGTCAACCACTAGTCAATAGCCATGAGCAACAGCAATGGAATGAAA 1335	
DB	601 ATTCTCATGCAAGTCAACCACTAGTCAATAGCCATGAGCAACAGCAATGGAATGAAA 660	
QY	1336 GAGGACCACTTTTCAGTCACTCTCTCTCAAAACATAGAAGAAAGTGCCTATTGTTGATT 1395	
DB	661 GAGGACCACTTTTCAGTCACTCTCTCTCAAAACATAGAAGAAAGTGCCTATTGTTGATT 720	
QY	1396 CCAGCTGGAAAGGCTTAACAGCTCTAGGAGGCTGTATTTTCATGTTCTTGTGTAACATG 1455	
DB	721 CCAGCTGGAAAGGCTTAACAGCTCTAGGAGGCTGTATTTTCATGTTCTTGTGTAACATG 780	
QY	1456 TCCTCAGATTGATCAACAAATTTAAAGATAGAAGAAAGAAATCAGAAGAAACCTGAAA 1515	
DB	781 TCCTCAGATTGATCAACAAATTTAAAGATAGAAGAAAGAAATCAGAAGAAACCTGAAA 840	
QY	1516 ATGATGATGATGTGGAGATTAAAGAGCAGTGTGTCCAAGTATGAATCTCAACTTTCAACAA 1575	
DB	841 ATGATGATGATGTGGAGATTAAAGAGCAGTGTGTCCAAGTATGAATCTCAACTTTCAACAA 900	
QY	1576 ATGAGGAGAAAGTATAGATACAGATGATGATGAGGCTTATTTACGAGCAGACTCAACAG 1635	
DB	901 ATGAGGAGAAAGTATAGATACAGATGATGATGAGGCTTATTTACGAGCAGACTCAACAG 960	
QY	1636 AGCCCTCCCACTTTGATTTCTCAGAGGCTGTGAGTCTTGGAGAGAAAGAGTCTATGATAG 1695	
DB	961 AGCCCTCCCACTTTGATTTCTCAGAGGCTGTGAGTCTTGGAGAGAAAGAGTCTATGATAG 1020	
QY	1696 CTCATGCTCATCCACAGGAAGTCTACATGAATATGTACCCAGAGGGTGCAGAAATAAAT 1755	
DB	1021 CTCATGCTCATCCACAGGAAGTCTACATGAATATGTACCCAGAGGGTGCAGAAATAAAT 1080	
QY	1756 GCCATTACATTTCCACGATACACTCGGCCAGTCCAGAGCTCTCATTTACACCACCATCATG 1815	
DB	1081 GCCATTACATTTCCACGATACACTCGGCCAGTCCAGAGCTCTCATTTACACCACCATCATG 1140	
QY	1816 ACT 1818	
DB	1141 ACT 1143	
RESULT 12		
AAK79819		
ID	AAK79819	standard; DNA; 20778 BP.
XX	XX	
AC	AAK79819;	
XX	XX	
DT	07-NOV-2001	(first entry)
XX	XX	
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34631.	
XX	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;	
KW		

Tue Nov 26 12:43:33 2002

us-09-642-034-1.l.rng

cytostatic; gene therapy; vaccine; metastasis; ds.

homo sapiens.

WO200157182-A2.

09-AUG-2001.

17-JAN-2001; 2001WO-US01354.

31-JAN-2000; 2000US-0179065.

04-FEB-2000; 2000US-0180628.

24-FEB-2000; 2000US-0184664.

02-MAR-2000; 2000US-0186350.

16-MAR-2000; 2000US-0189874.

17-MAR-2000; 2000US-0190076.

18-APR-2000; 2000US-0198123.

19-MAY-2000; 2000US-0205515.

07-JUN-2000; 2000US-0209467.

28-JUN-2000; 2000US-0214886.

30-JUN-2000; 2000US-0215135.

07-JUL-2000; 2000US-0216647.

07-JUL-2000; 2000US-0216880.

11-JUL-2000; 2000US-0217496.

11-JUL-2000; 2000US-0218290.

14-JUL-2000; 2000US-0220963.

26-JUL-2000; 2000US-0220964.

26-JUL-2000; 2000US-0224518.

14-AUG-2000; 2000US-0224519.

14-AUG-2000; 2000US-0225213.

14-AUG-2000; 2000US-0225214.

14-AUG-2000; 2000US-0225266.

14-AUG-2000; 2000US-0225267.

14-AUG-2000; 2000US-0225268.

14-AUG-2000; 2000US-0225270.

14-AUG-2000; 2000US-0225447.

14-AUG-2000; 2000US-0225757.

14-AUG-2000; 2000US-0225758.

14-AUG-2000; 2000US-0225759.

18-AUG-2000; 2000US-0226279.

22-AUG-2000; 2000US-0226681.

22-AUG-2000; 2000US-0226868.

22-AUG-2000; 2000US-0227182.

23-AUG-2000; 2000US-0227009.

30-AUG-2000; 2000US-0228924.

01-SEP-2000; 2000US-0229287.

01-SEP-2000; 2000US-0229343.

01-SEP-2000; 2000US-0229344.

01-SEP-2000; 2000US-0229345.

05-SEP-2000; 2000US-0229509.

05-SEP-2000; 2000US-0229513.

06-SEP-2000; 2000US-0230437.

06-SEP-2000; 2000US-0230438.

08-SEP-2000; 2000US-0231242.

08-SEP-2000; 2000US-0231243.

08-SEP-2000; 2000US-0231244.

08-SEP-2000; 2000US-0231413.

08-SEP-2000; 2000US-0231414.

08-SEP-2000; 2000US-0232080.

08-SEP-2000; 2000US-0232081.

12-SEP-2000; 2000US-0231968.

14-SEP-2000; 2000US-0232397.

14-SEP-2000; 2000US-0232398.

14-SEP-2000; 2000US-0232399.

14-SEP-2000; 2000US-0232400.

14-SEP-2000; 2000US-0232401.

14-SEP-2000; 2000US-0233063.

14-SEP-2000; 2000US-0233064.

14-SEP-2000; 2000US-0233065.

21-SEP-2000; 2000US-0234223.

21-SEP-2000; 2000US-0234274.

25-SEP-2000; 2000US-0234997.

25-SEP-2000; 2000US-0234998.

26-SEP-2000; 2000US-0235484.

27-SEP-2000; 2000US-0235834.

27-SEP-2000; 2000US-0235836.

29-SEP-2000; 2000US-0236327.

29-SEP-2000; 2000US-0236367.

29-SEP-2000; 2000US-0236368.

29-SEP-2000; 2000US-0236369.

29-SEP-2000; 2000US-0236370.

02-OCT-2000; 2000US-0236802.

02-OCT-2000; 2000US-0237037.

02-OCT-2000; 2000US-0237038.

02-OCT-2000; 2000US-0237039.

13-OCT-2000; 2000US-0237040.

13-OCT-2000; 2000US-0239935.

13-OCT-2000; 2000US-0239937.

20-OCT-2000; 2000US-0240960.

20-OCT-2000; 2000US-0241221.

20-OCT-2000; 2000US-0241785.

20-OCT-2000; 2000US-0241786.

20-OCT-2000; 2000US-0241787.

20-OCT-2000; 2000US-0241808.

20-OCT-2000; 2000US-0241809.

20-OCT-2000; 2000US-0241826.

01-NOV-2000; 2000US-0244617.

08-NOV-2000; 2000US-0246474.

08-NOV-2000; 2000US-0246475.

08-NOV-2000; 2000US-0246476.

08-NOV-2000; 2000US-0246477.

08-NOV-2000; 2000US-0246478.

08-NOV-2000; 2000US-0246523.

08-NOV-2000; 2000US-0246524.

08-NOV-2000; 2000US-0246525.

08-NOV-2000; 2000US-0246526.

08-NOV-2000; 2000US-0246527.

08-NOV-2000; 2000US-0246528.

08-NOV-2000; 2000US-0246532.

08-NOV-2000; 2000US-0246609.

08-NOV-2000; 2000US-0246610.

08-NOV-2000; 2000US-0246611.

08-NOV-2000; 2000US-0246613.

17-NOV-2000; 2000US-0249207.

17-NOV-2000; 2000US-0249208.

17-NOV-2000; 2000US-0249209.

17-NOV-2000; 2000US-0249210.

17-NOV-2000; 2000US-0249211.

17-NOV-2000; 2000US-0249212.

17-NOV-2000; 2000US-0249213.

17-NOV-2000; 2000US-0249214.

17-NOV-2000; 2000US-0249215.

17-NOV-2000; 2000US-0249216.

17-NOV-2000; 2000US-0249217.

17-NOV-2000; 2000US-0249218.

17-NOV-2000; 2000US-0249244.

17-NOV-2000; 2000US-0249245.

17-NOV-2000; 2000US-0249264.

17-NOV-2000; 2000US-0249265.

17-NOV-2000; 2000US-0249287.

17-NOV-2000; 2000US-0249289.

17-NOV-2000; 2000US-0249299.

01-DEC-2000; 2000US-0250160.

01-DEC-2000; 2000US-0250391.

05-DEC-2000; 2000US-0251030.

05-DEC-2000; 2000US-0251988.

05-DEC-2000; 2000US-0256719.

06-DEC-2000; 2000US-0251479.

08-DEC-2000; 2000US-0251856.

08-DEC-2000; 2000US-0251868.

08-DEC-2000; 2000US-0251869.

08-DEC-2000; 2000US-0251989.

11-DEC-2000; 2000US-0251990.

05-JAN-2001; 2001US-0259678.


```
FT /transl_except= (pos:879..881,aa:Xaa)
FT /note= "Xaa = unknown; the CDS does not include a
FT stop codon"
FT
FT unsure 535..537
FT /*tag= b
FT /note= "encodes Lys"
FT 538..540
FT /*tag= c
FT /note= "encodes Lys"
FT
XX WO200216939-A2.
XX
XX 28-FEB-2002.
XX
XX 20-AUG-2001: 2001WO-US25997.
XX
XX 18-AUG-2000; 2000US-0642034.
XX 08-DEC-2000; 2000US-0733320.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Mack D, Gish KC, Wilson KE;
XX
XX WPI; 2002-242033/29.
XX P-PSDB; AAM51199.
XX
XX Screening for drugs that affect expression of the breast cancer 4 gene
XX or its fragments, use of these to treat prostate and breast cancer, and
XX diagnosing these diseases -
XX
XX Disclosure; Fig 3; 83pp; English.
XX
XX The presence sequence is that of a partial cDNA for the murine
XX breast cancer 4 gene (BCR4). The cDNA encodes the protein given
XX in AAM51199. The invention also provides human BCR4 cDNA (see
XX AAM9299) and protein (see AAM51198) sequences. Methods are claimed
XX for using BCR4 polynucleotides, polypeptides and antibodies to:
XX screen drug candidates; screen bioactive agents capable of binding
XX to BCR4; evaluate the effect of a prostate cancer and/or breast
XX cancer drug; diagnose prostate cancer or breast cancer; screen for
XX a bioactive agent capable of interfering with the binding of BCR4 to
XX an antibody which binds BCR4; neutralise the effect of BCR4; treat
XX prostate cancer or breast cancer using an inhibitor of BCR4; treat
XX an antibody; localise a therapeutic moiety to, or treat, prostate
XX cancer or breast cancer tissue by conjugating an antibody for BCR4
XX to a cytotoxic agent or radioisotope; inhibit prostate cancer or
XX breast cancer by administering an antisense molecule to the present
XX nucleic acid sequence; elicit an immune response by administering
XX BCR4 or a nucleic acid encoding BCR4; and determine the prognosis
XX of an individual with prostate cancer or breast cancer by
XX determining the level of BCR4 in a sample, a high level indicating
XX a poor prognosis. A biochip comprising BCR4 nucleic acids is also
XX claimed.
XX
XX Sequence 899 BP; 252 A; 239 C; 198 G; 203 T; 7 other;
XX
XX
XX Query Match 17.9%; Score 619.6; DB 24; Length 899;
XX Best Local Similarity 84.3%; Pred. No. 4.8e-144;
XX Matches 752; Conservative 1; Mismatches 131; Indels 8; Gaps 5;
XX
XX QY 960 GGCATGGGATCCAGGTTCCGCTGAATGCAACAGAGTTCAACTATCTCTGTCAGCCATC 1019
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 1 GGCATGGGATCCAGGTTCCGCTGAATGCAACAGAGTTCAACTATCTCTGTCAGCCATC 60
XX
XX QY 1020 ATCAACCAATTCATGCTAGATCTTCTGCTGATTATATA---CAAGTGAAAGAGGCTGAA 1076
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 61 ATCAACCAATTCATGCTAGGCTTCTGCTGATTATCATACAGCAAGTGAGAAGAGGAGAA 120
XX
XX QY 1077 ATCCCTCCCAAGACCTATTATATACAAATAGGCTGGGTGGTGGTTTATAGCCATTTC 1136
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 121 ATCCCTCCCAAGACCTATTCTTTACAAATAGGCTGGGTGGTGGTTCATAGCCATTTC 180
XX
XX QY 1137 ATCATCAGTTTCTGCTGCTGGGGTTATCTTAGTGCCCTCATGAATCGGGTGT 1196
```

```
Db 181 ATCATCAGTTTCTGCTGCTGGAGTCATCTTGGTGCCACATCATGAACGGGTATTT 240
QY 1197 TTCAAAATTTCTCTGAGTTTCTGCTGGAGTGGCGCTTGGAGCTTTGAGTGGTGTGCT 1256
Db 241 TTCAAGTTCTCTGCTGAGCTTCTGCTGGCGCTGGGAGCTGCGGAGCTGAGTGGGATGCT 300
QY 1257 TTTTACACCTTTCTTCCACATTTCTCATCAAGTCAACCACTAGTATAGGCATGAAGAA 1316
Db 301 CTGTTACATCTTCTCCACACTCTCATCAAGTCAACCACTAGTATAGGCATGAAGAG 360
QY 1317 CCAGCAATGGAATGAAAGAGGAGGACCACTTTTTCAGTCACTCTCTCTCAAAACATAGAA 1376
Db 361 CCAGCGATGGAATGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 1377 GAAAGTGCCTATTTTGAATTCACAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1436
Db 421 GAAAGTGCCTATTTTGAATTCACAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 1437 ATGTTTCTTGTGAAACATGCTCTCATCAATGATGATGATGATGATGATGATGATGATGAT 1496
Db 481 ATGTTTCTTGTGAAACATGCTCTCATCAATGATGATGATGATGATGATGATGATGATGAT 540
QY 1497 AATCAGCAAGAAACCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1556
Db 541 AANMAAAGAAACCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 1557 GAATCTCAACTTTTCAACAAATGAGGAGAAAGTAGATACAGATGATGATGATGATGATGAT 1616
Db 601 GACTCTCAGCTTTCTTCAATGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 660
QY 1617 TTACGAGCAGACTCACAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1675
Db 661 CTGCGAGCAGACTCACAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
QY 1676 AGAAGAGAGGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1732
Db 721 AGAAGAGAGGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 1733 ACCCAGAGGAGTGAAGAAATGAATGCCATTCACATTTCCAGATACACTCGGCGAGTCAGA 1792
Db 781 CCCAGGGGGCTGCAAGAAACAAGTGCCATTCACACTTTTCAGATACGCTGGGCGAGTCCGA 840
QY 1793 CGATCTCATTTTCAACCACTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1844
Db 841 CGACCT-ATNACACACCATCAGACTACCATCATCATCTTCTGNACCACCCCTC 891
XX
XX RESULT 14
XX AAD13487
XX ID AAD13487 standard; cDNA; 551 BP.
XX AC AAD13487;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human LIV-1-164647 cDNA fragment (nucleotides 1690-2240).
XX
XX Human; LIV-1-164647; cytostatic; estrogen-inducible gene; tumour;
XX cancer; breast; lung; prostate; colon; ovary; uterus; kidney; gastric;
XX salivary gland; carcinoma; drug screening; therapy; ss.
XX
XX Homo sapiens.
XX
XX WO200155178-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US02622.
XX
XX 25-JAN-2000; 2000US-0177951.
XX 10-APR-2000; 2000US-0195761.
XX
```

PA (GETH) GENENTECH INC.
XX
XX
PI Goddard A, Gurney AL, Smith V, Hongo JS, De Sauvage F;
XX
XX WPI: 2001-502628/55.
XX
XX New LIV-1 DNA and amino acids, useful for treating cancer or inhibiting
PT the proliferation of tumor cells in mammals, e.g. breast, lung,
PT prostate, colon, ovary, uterus, kidney, gastric or salivary carcinoma
PT .
XX
XX
XX Example 5; Page 76-77; 150pp; English.
PS
PS The present sequence is human estrogen-inducible LIV-1-164647 cDNA
CC fragment (nucleotides 1690-2240). LIV-1 is overexpressed in tumor
CC tissues such as prostate, colon, lung and breast. The LIV-1 DNA and
CC polypeptide, and the monoclonal antibody are useful for treating cancer
CC and inhibiting the proliferation of tumor cells in mammals, e.g. breast,
CC lung, prostate, colon, ovary, uterus, kidney, gastric or salivary
CC carcinoma, or other tumor cell types expressing the LIV-1-164647
CC protein. In particular, the mammal is a human. The LIV-1 DNA and
CC polypeptide may also be used in screening assays for drug candidates.
XX
XX Sequence 551 BP; 128 A; 130 C; 127 G; 166 T; 0 other;
SQ
Query Match 15.9%; Score 549.4; DB 22; Length 551;
Best Local Similarity 99.8%; Pred. No. 1.2e-126;
Matches 550; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1755 TGCCATTACATTTCCAGATACACTCGCGCAGTCAGACGATCTCATTCACACCATCAT 1814
DB 1 TGCCATTACATTTCCAGATACACTCGCGCAGTCAGACGATCTCATTCACACCATCAT 60
QY 1815 GACTACCATCATATTTCCATCATCACCACCAACCAACCAATCCATCAGTCACAGC 1874
DB 61 GACTACCATCATATTTCCATCATCACCACCAACCAACCAATCCATCAGTCACAGC 120
QY 1875 CAGCGCTACTCTCGGAGAGCTGAAGTGCAGCGCTGCACATTTGGCTGGATGCTG 1934
DB 121 CAGCGCTACTCTCGGAGAGCTGAAGTGCAGCGCTGCACATTTGGCTGGATGCTG 180
QY 1935 ATAAATGGGTGATGGCTGCACAAATTCAGCGATGGCCTAGCAATTTGGTCTTTTACT 1994
DB 181 ATAAATGGGTGATGGCTGCACAAATTCAGCGATGGCCTAGCAATTTGGTCTTTTACT 240
QY 1995 GAAGGCTTATCAAGTGGTTTAAGTACTTCTGTCTGTCTTCTGTCTGATGCTGCTCAT 2054
DB 241 GAAGGCTTATCAAGTGGTTTAAGTACTTCTGTCTGTCTTCTGTCTGATGCTGCTCAT 300
QY 2055 GAATTAGGTGACTTTTCTGTCTTCTTAAAGGCTGGCATGACCGTTAAGCAGGCTGCTCT 2114
DB 301 GAATTAGGTGACTTTTCTGTCTTCTTAAAGGCTGGCATGACCGTTAAGCAGGCTGCTCT 360
QY 2115 TATAATGCATTTGACGCCATGCTGGCGTATCTTGGAAATGGCAACAGGAATTTTCATGCT 2174
DB 361 TATAATGCATTTGACGCCATGCTGGCGTATCTTGGAAATGGCAACAGGAATTTTCATGCT 420
QY 2175 CATTTAGCTGAAATGTTTCTATGTGATATTTGCACTTACTGCTGGCTTATTCATGAT 2234
DB 421 CATTTAGCTGAAATGTTTCTATGTGATATTTGCACTTACTGCTGGCTTATTCATGAT 480
QY 2235 GTTGCTCTGCTGATGATGTTGATGCTGAAATGCTGCACAAATGATGCTAGTGACCATGATG 2294
DB 481 GTTGCTCTGCTGATGATGTTGATGCTGAAATGCTGCACAAATGATGCTAGTGACCATGATG 540
QY 2295 AGCCGCTGGGG 2305
DB 541 AGCCGCTGGGG 551
RESULT 15
ABL79043
ID ABL79043 standard; cDNA; 483 BP.

XX ABL79043;
AC
XX 17-MAY-2002 (first entry)
DT
XX Human ovarian cancer related cDNA clone SEQ ID NO:2021.
DE
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
KW
XX Homo sapiens.
OS
XX WO200192581-A2.
PN
XX 06-DEC-2001.
PD
XX 29-MAY-2001; 2001WO-US17756.
PF
XX 26-MAY-2000; 2000US-207484P.
PR
XX (CORI-) CORIXA CORP.
PA
XX Algate PA, Harlocker SL, Jones R;
PI
XX WPI: 2002-122075/16.
DR
XX
XX Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide .
XX
XX Claim 1; SEQ ID 2021; 489pp; English.
PS
XX The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
CC population of (II), or antigen presenting cells that express (II).
CC (S1) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (S1) can be used for detecting ovarian cancer in a patient's biological
CC sample preferably serum or ovarian tissue. The method comprises
CC contacting a biological sample from a patient with (IV), detecting the
CC amount of polynucleotide hybridising to (IV) and comparing the amount to
CC a predetermined cutoff value and thereby detecting ovarian cancer in the
CC patient, where the amount of polynucleotide hybridising to (IV) is
CC detected preferably by polymerase chain reaction (PCR). (I) comprising
CC (III) and/or (II) is useful for stimulating and/or expanding T cells
CC specific for an ovarian tumour protein comprising contacting T cells
CC with (III) or (II). (III) is useful in design and preparation of
CC ribozyme molecules for inhibiting expression of the tumour polypeptides
CC and proteins in tumour cells; and to isolate a full length gene from a
CC suitable library e.g., a tumour cDNA library using well known
CC techniques.
XX
SQ Sequence 483 BP; 165 A; 61 C; 99 G; 150 T; 8 other;
Query Match 11.7%; Score 404.8; DB 24; Length 483;
Best Local Similarity 92.1%; Pred. No. 1.2e-90;
Matches 444; Conservative 0; Mismatches 35; Indels 3; Gaps 2;
QY 2541 TGTTCACAAAGTCAGTTAAAGGTACGTTTAAATATTAGTTATTCTATCTTGGAGATAA 2600
DB 2 TGTTCACAAAGTCAGTTAAAGGTACGTTTAAATATTAGTTATTCTATCTTGGAGATAA 61
QY 2601 ATCTGTATGTGCAATTCACCGGTATTACAGTTTATTATGTAACAAGAGATTGGCATG 2660
DB 62 ATCTGTATGTGCAATTCACCGGTATTACAGTTTATTATGTAACAAGAGATTGGCATG 121
QY 2661 ACATGTTCTGTATGTTTCAGGAAAAATGCTTTAATGCTTTTCAACAACATACACAGT 2720
DB 122 ACATGTTCTGTATGTTTCAGGAAAAATGCTTTAATGCTTTTCAACAACATACACAGT 181
QY 2721 TATTCCTATCTACTGGATTTTATGCTCTCTGAAGAACTGCTGGTGTGTAGGAATAAGAACTG 2780

Tue Nov 26 12:43:33 2002

Db	182	TATTCCTATAC	TTTTAGGTCCTGAAGAACTGCTGGTGT	TTAGGAATAAGAAATCTG	241
Oy	2781	CATGAAGCCTAA	AAATACCAAGAAAGCTTATAC	TGAAATTTAAGCAAGAAATTAAGGAGAA	2840
Db	242	CATGAAGCCTAA	ATACCAAGAAAGCTTATAC	TGAAATTTAAGCAAGAAATTAAGGAGAA	301
Oy	2841	AAGAGAAGAATC	--TGAGAATTGGGGAGGCATAGATTC	TATATAAAATCACAATAATTTGT	2898
Db	302	AAGAGAAGAATC	CTGAGAATTGGGGAGGCATAGATTC	TATATAAAATCACAATAATTTGT	361
Oy	2899	TGTAAT	TTAGAGGGAGAAATTTAGAAATTAAGTATATAAAAGGCAGAAATTAAGTATAGAT	2957	
Db	362	TGTAAT	TTAGAGGGAGAAATTTAGAAATTAAGTATATAAAAGGCAGAAATTAAGTATAGAT	421	
Oy	2958	ACATTCATTAACA	ATTTTGTGAGGATTTTCCCGTAAACGTAAGTGAACACTCTCAT	3017	
Db	422	ACNTTCCNTTAACA	TTTTTGCCCGGATTTATCCCGTAAACGTAAGTGAACACTCTCAT	481	
Oy	3018	AT	3019		
Db	482	AT	483		

Search completed: November 8, 2002, 17:48:57
Job time : 603.31 secs

[illegible]

	GAACTAACACAGTATTTCCTACTGATTTTAGGTCTTGGAAGA	CCTGGTGTTAG	2702
DJ68	GAATAAAGCATGCATGAAGCCTTAATAACCAGAAAAGCTTAT	TACTGTAATTTAAACAAG	2827
DB	GAATAAAGCATGCATGAAGCCTTAATAACCAGAAAAGCTTAT	TACTGTAATTTAAACAAG	2762
OY	2828	AAAAAAGGAGAAA	2841
DB	2763	AAAAAAAAAAAAA	2776
RESULT #	HSU41060		
LOCUS	Homo sapiens estrogen regulated LIV-1 protein (LIV-1) mRNA,	2744 bp	PRI 08-FEB-2001
DEFINITION	complete cds.		
ACCESSION	U41060		
KEYWORDS	U41060.2 GI:12711792		
SOURCE	.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 2744)		
JOURNAL	Green,C. and Morgan,H. Direct Submission Submitted (08-FEB-2001) Biochemistry, University of Liverpool, P.O.		
REFERENCE	Box 147, Liverpool L69 3BX, UK		
AUTHORS	2 (bases 1 to 2744)		
TITLE	Green.C., Gilhooley,E.M. and Walker,N.J.		
JOURNAL	Direct Submission Submitted (21-Nov-1995) Biochemistry, University of Liverpool, P.O.		
COMMENT	Box 147, Liverpool L69 3BX, UK		
FEATURES	On Feb 8, 2001 this sequence version replaced gi:1256000.		
source	1..2744		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
gene	/cell_line="MCF-7 human breast cancer cell line"		
	/note="estrogen induced mRNA"		
	1..2744		
CDS	/gene="Liv-1"		
	138..2387		
	/gene="Liv-1"		
	/note="breast cancer"		
	/codon_start=1		
	/product="estrogen regulated Liv-1 protein"		
	/protein_id="AAA96258.2"		
	/db_xref="GI:12711793"		
	/translation="MARKLSVLTLLTFALSVTNPLNHELKKAAPQPTTEKISPWNWESI NVDLATSTROYHLQOLFERYKGNNSLSEVEGRKLQNIGIDIKRIRIHHDHSHD EHEDHERSDHEHDDHHASHGNKKRALCPDPHDSGSCKPDRNSQGKG HREHASGRHNVDSSVASEVTNTYTNTVSGETHFLETETPRPKLFKDYSVP SVTSKSRLVLAGRTNESVSPPRGPMYSRNTPENPOECFNASKLLTSGMGIOGPL NATFYELPAIINGITADRSCLHIHTSEKKAEIPKTYSLOIANVGGFIALISFELS LGVLVLMNRVFEFKFLSFVALAVGLSCDAIFMLHPHSHAHHSSHSEPAMEX KRCPLFSHLSNONIEESAVIDFTSKGTLALGGLFMELVELHVLTQLKFDKKNKNOX KPNDODDVLEIKOLSKYESQLSTNEEKVDTOOTREGYLRADESPESHQQPFALEE EFWIAHAHPQEVIINYVRGCKNKSHFHDTLGQSDDLIIHHHYDHILHSHHHON HHPSHSHREEKLADVAGTWIAMVMIMGDLNFDFSGDLAIQAATEGSLSGISTSY AVFCHELPELDGFAVLAKAGMTAQVAVLYNALSAMLAYLGMATGFIIGHYAENVSM IFALTAGFMLVALVDMPVELMLHNASDHGCSRWGYFFLNAGMLLFGFILMLLIISIFE HKIVFRINF"		
BASE COUNT	786 A	585 C	601 G
ORIGIN	772 T		
Query Match	78.5%	Score 2716;	DB 9; Length 2744;
Best Local Similarity	99.3%;	Pred. No. 0;	
Matches 2744:	Conservative	0; Mismatches	0; Indels 18; Gaps 1;
OY	1 CTCGCGCCGCAATTCGCCAGCACGACCGGTGTCGCCGCTGGTAGAGATTTCTCGAAGACA	60	

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 8, 2002, 16:42:31 ; Search time 6148.03 Seconds
(without alignments)
16383.279 Million cell updates/sec

Title: US-09-642-034-1
Perfect score: 3461
Sequence: 1 ctcgtccgaattccgacg.....aaattcttatatcagcttg 3461

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pt.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3461	100.0	3461	6	AX465588	AX465588 Sequence
2	3391	98.0	3461	6	AX207205	AX207205 Sequence
3	2761.2	79.8	2776	6	AX207207	AX207207 Sequence
4	2716	78.5	2744	9	HSU41060	U41060 Homo sapien
5	2249	65.0	2404	6	I76892	I76892 Sequence 3
6	2134.2	61.7	2236	9	BC008317	BC008317 Homo sapi
7	1753.2	50.7	3287	10	AB071697	AB071697 Homo sapi
8	1486.8	43.0	1597	6	AX017261	AX017261 Sequence
9	1288	37.2	1310	6	I76891	I76891 Sequence 1
10	1137.8	32.9	221941	9	AC091060	AC091060 Homo sapi
11	1137.8	32.9	224788	2	AF001905	AF001905 Homo sapi
12	804	23.2	160170	2	AF001158	AF001158 Homo sapi
13	619.6	17.9	899	6	AX465590	AX465590 Sequence
14	549.4	15.9	551	6	AX207216	AX207216 Sequence
15	429.2	12.4	193437	2	AC126667	AC126667 Rattus no
16	390.8	11.3	395	6	AX158237	AX158237 Sequence
17	336.8	9.7	445	11	G24009	G24009 human STS W
18	319.2	9.2	337	6	AX393861	AX393861 Sequence
19	319	9.2	5231	9	AB033091	AB033091 Homo sapi
20	319	9.2	5322	6	AX405756	AX405756 Sequence
21	282.8	8.2	61103	2	AC131276	AC131276 Homo sapi
22	253.8	7.3	391	6	AX207209	AX207209 Sequence
23	240	6.9	241	6	AX303040	AX303040 Sequence
24	191.2	5.5	193437	2	AC126667	AC126667 Rattus no
25	165	4.8	177	6	AX106381	AX106381 Sequence
26	165	4.8	177	6	AX140672	AX140672 Sequence
27	165	4.8	177	6	AX200532	AX200532 Sequence
28	165	4.8	177	6	AX267188	AX267188 Sequence
29	144	4.2	52684	2	AC014796	AC014796 Drosophil
30	144	4.2	174712	3	AC010041	AC010041 Drosophil
31	144	4.2	195148	3	AC010019	AC010019 Drosophil
32	144	4.2	282267	3	AE003555	AE003555 Drosophil
33	129.2	3.7	2174	10	BC021530	BC021530 Mus muscu
34	128.4	3.7	2015	9	AB051127	AB051127 Macaca fa
35	127.6	3.7	4573	9	AX090002	AX090002 Sequence
36	127.6	3.7	4573	9	HUMORFKG1P	D31887 Human mRNa
37	126	3.6	4610	6	AX078294	AX078294 Sequence
38	126	3.6	4698	6	AX281784	AX281784 Sequence
39	123.4	3.6	1383	6	AX118897	AX118897 Sequence
40	123.4	3.6	1803	9	AF193052	AF193052 Homo sapi
41	123.4	3.6	2604	9	AK027652	AK027652 Homo sapi
42	123.4	3.6	3169	9	BC012125	BC012125 Homo sapi
43	123.4	3.6	3246	9	AB040120	AB040120 Homo sapi
44	122.6	3.5	1895	9	BC027884	BC027884 Homo sapi
45	121.4	3.5	55577	2	AC014109	AC014109 Drosophil

ALIGNMENTS

RESULT 1
AX465588
LOCUS AX465588 3461 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO0216939.
ACCESSION AX465588
VERSION AX465588.1 GI:21899891
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Mack,D., Gish,K.C. and Wilson,K.E.
TITLE Methods of diagnosis of cancer and screening for cancer modulators
JOURNAL Patent: WO 0216939-A 1 28-FEB-2002;